

Deep Q-Learning

Install dependencies for AI gym to run properly (shouldn't take more than a minute).
If running on google cloud or running locally, only need to run once. Colab may require installing everytime the vm shuts down.

```
In [ ]: !pip3 install gym pyvirtualdisplay
        !sudo apt-get install -y xvfb python-opengl ffmpeg
```

```
Requirement already satisfied: gym in /home/wxzhang/anaconda3/envs/cv/lib/
python3.8/site-packages (0.26.2)
Requirement already satisfied: pyvirtualdisplay in /home/wxzhang/anaconda3
/envs/cv/lib/python3.8/site-packages (3.0)
Requirement already satisfied: numpy>=1.18.0 in /home/wxzhang/anaconda3/en
vs/cv/lib/python3.8/site-packages (from gym) (1.24.4)
Requirement already satisfied: cloudpickle>=1.2.0 in /home/wxzhang/anacond
a3/envs/cv/lib/python3.8/site-packages (from gym) (3.0.0)
Requirement already satisfied: gym-notices>=0.0.4 in /home/wxzhang/anacond
a3/envs/cv/lib/python3.8/site-packages (from gym) (0.0.8)
Requirement already satisfied: importlib-metadata>=4.8.0 in /home/wxzhang/
anaconda3/envs/cv/lib/python3.8/site-packages (from gym) (7.1.0)
Requirement already satisfied: zipp>=0.5 in /home/wxzhang/anaconda3/envs/c
v/lib/python3.8/site-packages (from importlib-metadata>=4.8.0->gym) (3.18.
1)
[sudo] password for wxzhang:
```

```
In [ ]: !pip3 install --upgrade setuptools --user
        !pip3 install ez_setup
        !pip3 install gym[atari]
        !pip3 install gym[accept-rom-license]
```

```
Requirement already satisfied: setuptools in /home/wxzhang/.local/lib/pyth
on3.8/site-packages (69.5.1)
Requirement already satisfied: ez_setup in /home/wxzhang/anaconda3/envs/cv
/lib/python3.8/site-packages (0.9)
Requirement already satisfied: gym[atari] in /home/wxzhang/anaconda3/envs/
cv/lib/python3.8/site-packages (0.26.2)
Requirement already satisfied: numpy>=1.18.0 in /home/wxzhang/anaconda3/en
vs/cv/lib/python3.8/site-packages (from gym[atari]) (1.24.4)
Requirement already satisfied: cloudpickle>=1.2.0 in /home/wxzhang/anacond
a3/envs/cv/lib/python3.8/site-packages (from gym[atari]) (3.0.0)
Requirement already satisfied: gym-notices>=0.0.4 in /home/wxzhang/anacond
a3/envs/cv/lib/python3.8/site-packages (from gym[atari]) (0.0.8)
Requirement already satisfied: importlib-metadata>=4.8.0 in /home/wxzhang/
anaconda3/envs/cv/lib/python3.8/site-packages (from gym[atari]) (7.1.0)
Requirement already satisfied: ale-py~0.8.0 in /home/wxzhang/anaconda3/en
vs/cv/lib/python3.8/site-packages (from gym[atari]) (0.8.1)
Requirement already satisfied: importlib-resources in /home/wxzhang/anacon
da3/envs/cv/lib/python3.8/site-packages (from ale-py~0.8.0->gym[atari]) (
6.4.0)
```

Requirement already satisfied: typing-extensions in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from ale-py~=0.8.0->gym[atari]) (4.10.0)

Requirement already satisfied: zipp>=0.5 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from importlib-metadata>=4.8.0->gym[atari]) (3.18.1)

Requirement already satisfied: gym[accept-rom-license] in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (0.26.2)

Requirement already satisfied: numpy>=1.18.0 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from gym[accept-rom-license]) (1.24.4)

Requirement already satisfied: cloudpickle>=1.2.0 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from gym[accept-rom-license]) (3.0.0)

Requirement already satisfied: gym-notices>=0.0.4 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from gym[accept-rom-license]) (0.0.8)

Requirement already satisfied: importlib-metadata>=4.8.0 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from gym[accept-rom-license]) (7.1.0)

Requirement already satisfied: autorom~=0.4.2 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[accept-rom-license]) (0.4.2)

Requirement already satisfied: click in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[accept-rom-license]) (8.1.7)

Requirement already satisfied: requests in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[accept-rom-license]) (2.31.0)

Requirement already satisfied: tqdm in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[accept-rom-license]) (4.66.2)

Requirement already satisfied: importlib-resources in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[accept-rom-license]) (6.4.0)

Requirement already satisfied: AutoROM.accept-rom-license in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[accept-rom-license]) (0.6.1)

Requirement already satisfied: zipp>=0.5 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from importlib-metadata>=4.8.0->gym[accept-rom-license]) (3.18.1)

Requirement already satisfied: charset-normalizer<4,>=2 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from requests->autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[accept-rom-license]) (3.3.2)

Requirement already satisfied: idna<4,>=2.5 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from requests->autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[accept-rom-license]) (3.7)

Requirement already satisfied: urllib3<3,>=1.21.1 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from requests->autorom~=0.4.2->aut

```

rom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[accept-rom-license]) (2.2.1)
Requirement already satisfied: certifi>=2017.4.17 in /home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages (from requests->autorom~=0.4.2->autorom[accept-rom-license]~=0.4.2; extra == "accept-rom-license"->gym[accept-rom-license]) (2024.2.2)

```

For this assignment we will implement the Deep Q-Learning algorithm with Experience Replay as described in breakthrough paper "**Playing Atari with Deep Reinforcement Learning**". We will train an agent to play the famous game of **Breakout**.

```

In [ ]: %matplotlib inline

import sys
import gym
import torch
import pylab
import random
import numpy as np
from collections import deque
from datetime import datetime
from copy import deepcopy
import torch.nn as nn
import torch.optim as optim
import torch.nn.functional as F
from torch.autograd import Variable
from utils import find_max_lives, check_live, get_frame, get_init_state
from model import DQN, DQN_LSTM
from config import *

import matplotlib.pyplot as plt
# %load_ext autoreload
# %autoreload 2

```

Understanding the environment

In the following cell, we initialize our game of **Breakout** and you can see how the environment looks like. For further documentation of the of the environment refer to <https://gym.openai.com/envs>.

In breakout, we will use 3 actions "fire", "left", and "right". "fire" is only used to reset the game when a life is lost, "left" moves the agent left and "right" moves the agent right.

```

In [ ]: env = gym.make('BreakoutDeterministic-v4')
state = env.reset()

```

A.L.E: Arcade Learning Environment (version 0.8.1+53f58b7)
[Powered by Stella]

```
In [ ]: number_lives = find_max_lives(env)
        state_size = env.observation_space.shape
        action_size = 3 #fire, left, and right
```

```
/home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages/gym/utils/passive_env_checker.py:233: DeprecationWarning: `np.bool8` is a deprecated alias for `np.bool_`. (Deprecated NumPy 1.24)
    if not isinstance(terminated, (bool, np.bool8)):
```

Creating a DQN Agent

Here we create a DQN Agent. This agent is defined in the **agent.py**. The corresponding neural network is defined in the **model.py**. Once you've created a working DQN agent, use the code in agent.py to create a double DQN agent in **agent_double.py**. Set the flag "double_dqn" to True to train the double DQN agent.

Evaluation Reward : The average reward received in the past 100 episodes/games.

Frame : Number of frames processed in total.

Memory Size : The current size of the replay memory.

```
In [ ]: double_dqn = True # set to True if using double DQN agent

if double_dqn:
    from agent_double import Agent
else:
    from agent import Agent

agent = Agent(action_size)
evaluation_reward = deque(maxlen=evaluation_reward_length)
frame = 0
memory_size = 0
```

Main Training Loop

In this training loop, we do not render the screen because it slows down training significantly. To watch the agent play the game, run the code in next section "Visualize Agent Performance"

```
In [ ]: device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")

rewards, episodes = [], []
best_eval_reward = 0
for e in range(EPISODES):
```

```

done = False
score = 0

history = np.zeros([5, 84, 84], dtype=np.uint8)
step = 0
state, _ = env.reset()
next_state = state
life = number_lives

get_init_state(history, state, HISTORY_SIZE)

while not done:
    step += 1
    frame += 1

    # Perform a fire action if ball is no longer on screen to continue
    if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
        action = torch.tensor([[0]]).cuda()
    else:
        action = agent.get_action(np.float32(history[:4, :, :]) / 255)
    state = next_state
    next_state, reward, terminated, truncated, info = env.step(action)
    done = terminated or truncated

    frame_next_state = get_frame(next_state)
    history[4, :, :] = frame_next_state
    terminal_state = check_live(life, info['lives'])

    life = info['lives']
    r = reward

    # Store the transition in memory
    agent.memory.push(deepcopy(frame_next_state), action.cpu(), r, terminal_state)
    # Start training after random sample generation
    if (frame >= train_frame): # You can set train_frame to a lower value
        agent.train_policy_net(frame)
        # Update the target network only for Double DQN only
        if double_dqn and (frame % update_target_network_frequency) == 0:
            agent.update_target_net()
    score += reward
    history[:4, :, :] = history[1:, :, :]

    if done:
        evaluation_reward.append(score)
        rewards.append(np.mean(evaluation_reward))
        episodes.append(e)
        pylab.plot(episodes, rewards, 'b')
        pylab.xlabel('Episodes')
        pylab.ylabel('Rewards')
        pylab.title('Episodes vs Reward')
        pylab.savefig("./save_graph/breakout_dqn.png") # save graph for training

    # every episode, plot the play time

```

```

print("episode:", e, "  score:", score, "  memory length:",
      len(agent.memory), "  epsilon:", agent.epsilon, "  steps:",
      "  lr:", agent.optimizer.param_groups[0]['lr'], "  evaluation reward:",
      np.mean(evaluation_reward))

# if the mean of scores of last 100 episode is bigger than 5
### Change this save condition to whatever you prefer ###
if np.mean(evaluation_reward) > 6 and np.mean(evaluation_reward) > 1.5:
    save_dir = "./save_model/breakout_dqn_"+str(e)+".pth"
    torch.save(agent.policy_net, save_dir)
    best_eval_reward = np.mean(evaluation_reward)

```

```

episode: 0    score: 0.0    memory length: 123    epsilon: 1.0    steps: 123
lr: 0.0001    evaluation reward: 0.0
episode: 1    score: 2.0    memory length: 320    epsilon: 1.0    steps: 197
lr: 0.0001    evaluation reward: 1.0
episode: 2    score: 3.0    memory length: 545    epsilon: 1.0    steps: 225
lr: 0.0001    evaluation reward: 1.6666666666666667
episode: 3    score: 6.0    memory length: 910    epsilon: 1.0    steps: 365
lr: 0.0001    evaluation reward: 2.75
episode: 4    score: 0.0    memory length: 1032    epsilon: 1.0    steps: 122
lr: 0.0001    evaluation reward: 2.2
episode: 5    score: 1.0    memory length: 1202    epsilon: 1.0    steps: 170
lr: 0.0001    evaluation reward: 2.0
episode: 6    score: 0.0    memory length: 1325    epsilon: 1.0    steps: 123
lr: 0.0001    evaluation reward: 1.7142857142857142
episode: 7    score: 1.0    memory length: 1493    epsilon: 1.0    steps: 168
lr: 0.0001    evaluation reward: 1.625
episode: 8    score: 1.0    memory length: 1644    epsilon: 1.0    steps: 151
lr: 0.0001    evaluation reward: 1.5555555555555556
episode: 9    score: 3.0    memory length: 1875    epsilon: 1.0    steps: 231
lr: 0.0001    evaluation reward: 1.7
episode: 10   score: 2.0    memory length: 2076    epsilon: 1.0    steps: 20
1    lr: 0.0001    evaluation reward: 1.7272727272727273
episode: 11   score: 3.0    memory length: 2302    epsilon: 1.0    steps: 22
6    lr: 0.0001    evaluation reward: 1.8333333333333333
episode: 12   score: 0.0    memory length: 2425    epsilon: 1.0    steps: 12
3    lr: 0.0001    evaluation reward: 1.6923076923076923
episode: 13   score: 1.0    memory length: 2594    epsilon: 1.0    steps: 16
9    lr: 0.0001    evaluation reward: 1.6428571428571428
episode: 14   score: 2.0    memory length: 2813    epsilon: 1.0    steps: 21
9    lr: 0.0001    evaluation reward: 1.6666666666666667
episode: 15   score: 2.0    memory length: 2995    epsilon: 1.0    steps: 18
2    lr: 0.0001    evaluation reward: 1.6875
episode: 16   score: 0.0    memory length: 3118    epsilon: 1.0    steps: 12
3    lr: 0.0001    evaluation reward: 1.588235294117647
episode: 17   score: 1.0    memory length: 3287    epsilon: 1.0    steps: 16
9    lr: 0.0001    evaluation reward: 1.5555555555555556
episode: 18   score: 0.0    memory length: 3410    epsilon: 1.0    steps: 12
3    lr: 0.0001    evaluation reward: 1.4736842105263157
episode: 19   score: 2.0    memory length: 3607    epsilon: 1.0    steps: 19
7    lr: 0.0001    evaluation reward: 1.5
episode: 20   score: 3.0    memory length: 3852    epsilon: 1.0    steps: 24
5    lr: 0.0001    evaluation reward: 1.5714285714285714
episode: 21   score: 0.0    memory length: 3975    epsilon: 1.0    steps: 12

```

```
3   lr: 0.0001      evaluation reward: 1.5
episode: 22   score: 0.0   memory length: 4098   epsilon: 1.0   steps: 12
3   lr: 0.0001      evaluation reward: 1.434782608695652
episode: 23   score: 1.0   memory length: 4268   epsilon: 1.0   steps: 17
0   lr: 0.0001      evaluation reward: 1.4166666666666667
episode: 24   score: 3.0   memory length: 4515   epsilon: 1.0   steps: 24
7   lr: 0.0001      evaluation reward: 1.48
episode: 25   score: 1.0   memory length: 4684   epsilon: 1.0   steps: 16
9   lr: 0.0001      evaluation reward: 1.4615384615384615
episode: 26   score: 3.0   memory length: 4915   epsilon: 1.0   steps: 23
1   lr: 0.0001      evaluation reward: 1.5185185185185186
episode: 27   score: 0.0   memory length: 5038   epsilon: 1.0   steps: 12
3   lr: 0.0001      evaluation reward: 1.4642857142857142
episode: 28   score: 1.0   memory length: 5189   epsilon: 1.0   steps: 15
1   lr: 0.0001      evaluation reward: 1.4482758620689655
episode: 29   score: 2.0   memory length: 5387   epsilon: 1.0   steps: 19
8   lr: 0.0001      evaluation reward: 1.4666666666666666
episode: 30   score: 2.0   memory length: 5584   epsilon: 1.0   steps: 19
7   lr: 0.0001      evaluation reward: 1.4838709677419355
episode: 31   score: 0.0   memory length: 5706   epsilon: 1.0   steps: 12
2   lr: 0.0001      evaluation reward: 1.4375
episode: 32   score: 0.0   memory length: 5829   epsilon: 1.0   steps: 12
3   lr: 0.0001      evaluation reward: 1.3939393939393939
episode: 33   score: 0.0   memory length: 5952   epsilon: 1.0   steps: 12
3   lr: 0.0001      evaluation reward: 1.3529411764705883
episode: 34   score: 0.0   memory length: 6075   epsilon: 1.0   steps: 12
3   lr: 0.0001      evaluation reward: 1.3142857142857143
episode: 35   score: 2.0   memory length: 6273   epsilon: 1.0   steps: 19
8   lr: 0.0001      evaluation reward: 1.3333333333333333
episode: 36   score: 0.0   memory length: 6396   epsilon: 1.0   steps: 12
3   lr: 0.0001      evaluation reward: 1.2972972972972974
episode: 37   score: 2.0   memory length: 6593   epsilon: 1.0   steps: 19
7   lr: 0.0001      evaluation reward: 1.3157894736842106
episode: 38   score: 2.0   memory length: 6791   epsilon: 1.0   steps: 19
8   lr: 0.0001      evaluation reward: 1.3333333333333333
episode: 39   score: 3.0   memory length: 7037   epsilon: 1.0   steps: 24
6   lr: 0.0001      evaluation reward: 1.375
episode: 40   score: 2.0   memory length: 7220   epsilon: 1.0   steps: 18
3   lr: 0.0001      evaluation reward: 1.3902439024390243
episode: 41   score: 3.0   memory length: 7446   epsilon: 1.0   steps: 22
6   lr: 0.0001      evaluation reward: 1.4285714285714286
episode: 42   score: 1.0   memory length: 7615   epsilon: 1.0   steps: 16
9   lr: 0.0001      evaluation reward: 1.4186046511627908
episode: 43   score: 3.0   memory length: 7861   epsilon: 1.0   steps: 24
6   lr: 0.0001      evaluation reward: 1.4545454545454546
episode: 44   score: 1.0   memory length: 8033   epsilon: 1.0   steps: 17
2   lr: 0.0001      evaluation reward: 1.4444444444444444
episode: 45   score: 4.0   memory length: 8309   epsilon: 1.0   steps: 27
6   lr: 0.0001      evaluation reward: 1.5
episode: 46   score: 0.0   memory length: 8432   epsilon: 1.0   steps: 12
3   lr: 0.0001      evaluation reward: 1.4680851063829787
episode: 47   score: 3.0   memory length: 8679   epsilon: 1.0   steps: 24
7   lr: 0.0001      evaluation reward: 1.5
```

```
episode: 48  score: 2.0  memory length: 8877  epsilon: 1.0  steps: 19
8  lr: 0.0001  evaluation reward: 1.510204081632653
episode: 49  score: 2.0  memory length: 9098  epsilon: 1.0  steps: 22
1  lr: 0.0001  evaluation reward: 1.52
episode: 50  score: 1.0  memory length: 9270  epsilon: 1.0  steps: 17
2  lr: 0.0001  evaluation reward: 1.5098039215686274
episode: 51  score: 2.0  memory length: 9467  epsilon: 1.0  steps: 19
7  lr: 0.0001  evaluation reward: 1.5192307692307692
episode: 52  score: 1.0  memory length: 9639  epsilon: 1.0  steps: 17
2  lr: 0.0001  evaluation reward: 1.509433962264151
episode: 53  score: 3.0  memory length: 9850  epsilon: 1.0  steps: 21
1  lr: 0.0001  evaluation reward: 1.537037037037037
episode: 54  score: 0.0  memory length: 9973  epsilon: 1.0  steps: 12
3  lr: 0.0001  evaluation reward: 1.509090909090909
episode: 55  score: 2.0  memory length: 10152  epsilon: 1.0  steps: 1
79  lr: 0.0001  evaluation reward: 1.5178571428571428
episode: 56  score: 3.0  memory length: 10399  epsilon: 1.0  steps: 2
47  lr: 0.0001  evaluation reward: 1.543859649122807
episode: 57  score: 3.0  memory length: 10662  epsilon: 1.0  steps: 2
63  lr: 0.0001  evaluation reward: 1.5689655172413792
episode: 58  score: 7.0  memory length: 11048  epsilon: 1.0  steps: 3
86  lr: 0.0001  evaluation reward: 1.6610169491525424
episode: 59  score: 2.0  memory length: 11246  epsilon: 1.0  steps: 1
98  lr: 0.0001  evaluation reward: 1.6666666666666667
episode: 60  score: 1.0  memory length: 11416  epsilon: 1.0  steps: 1
70  lr: 0.0001  evaluation reward: 1.6557377049180328
episode: 61  score: 1.0  memory length: 11567  epsilon: 1.0  steps: 1
51  lr: 0.0001  evaluation reward: 1.6451612903225807
episode: 62  score: 0.0  memory length: 11690  epsilon: 1.0  steps: 1
23  lr: 0.0001  evaluation reward: 1.619047619047619
episode: 63  score: 3.0  memory length: 11938  epsilon: 1.0  steps: 2
48  lr: 0.0001  evaluation reward: 1.640625
episode: 64  score: 2.0  memory length: 12156  epsilon: 1.0  steps: 2
18  lr: 0.0001  evaluation reward: 1.646153846153846
episode: 65  score: 1.0  memory length: 12327  epsilon: 1.0  steps: 1
71  lr: 0.0001  evaluation reward: 1.6363636363636365
episode: 66  score: 1.0  memory length: 12498  epsilon: 1.0  steps: 1
71  lr: 0.0001  evaluation reward: 1.626865671641791
episode: 67  score: 2.0  memory length: 12718  epsilon: 1.0  steps: 2
20  lr: 0.0001  evaluation reward: 1.6323529411764706
episode: 68  score: 6.0  memory length: 13108  epsilon: 1.0  steps: 3
90  lr: 0.0001  evaluation reward: 1.6956521739130435
episode: 69  score: 0.0  memory length: 13231  epsilon: 1.0  steps: 1
23  lr: 0.0001  evaluation reward: 1.6714285714285715
episode: 70  score: 0.0  memory length: 13353  epsilon: 1.0  steps: 1
22  lr: 0.0001  evaluation reward: 1.647887323943662
episode: 71  score: 1.0  memory length: 13523  epsilon: 1.0  steps: 1
70  lr: 0.0001  evaluation reward: 1.6388888888888888
episode: 72  score: 2.0  memory length: 13720  epsilon: 1.0  steps: 1
97  lr: 0.0001  evaluation reward: 1.643835616438356
episode: 73  score: 4.0  memory length: 13999  epsilon: 1.0  steps: 2
79  lr: 0.0001  evaluation reward: 1.6756756756756757
episode: 74  score: 2.0  memory length: 14198  epsilon: 1.0  steps: 1
```



```
99   lr: 0.0001      evaluation reward: 1.68
episode: 75   score: 2.0   memory length: 14417   epsilon: 1.0   steps: 2
19   lr: 0.0001      evaluation reward: 1.6842105263157894
episode: 76   score: 0.0   memory length: 14540   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.6623376623376624
episode: 77   score: 1.0   memory length: 14709   epsilon: 1.0   steps: 1
69   lr: 0.0001      evaluation reward: 1.6538461538461537
episode: 78   score: 1.0   memory length: 14860   epsilon: 1.0   steps: 1
51   lr: 0.0001      evaluation reward: 1.6455696202531647
episode: 79   score: 1.0   memory length: 15030   epsilon: 1.0   steps: 1
70   lr: 0.0001      evaluation reward: 1.6375
episode: 80   score: 0.0   memory length: 15152   epsilon: 1.0   steps: 1
22   lr: 0.0001      evaluation reward: 1.617283950617284
episode: 81   score: 3.0   memory length: 15378   epsilon: 1.0   steps: 2
26   lr: 0.0001      evaluation reward: 1.6341463414634145
episode: 82   score: 1.0   memory length: 15547   epsilon: 1.0   steps: 1
69   lr: 0.0001      evaluation reward: 1.6265060240963856
episode: 83   score: 2.0   memory length: 15728   epsilon: 1.0   steps: 1
81   lr: 0.0001      evaluation reward: 1.630952380952381
episode: 84   score: 2.0   memory length: 15928   epsilon: 1.0   steps: 2
00   lr: 0.0001      evaluation reward: 1.6352941176470588
episode: 85   score: 3.0   memory length: 16157   epsilon: 1.0   steps: 2
29   lr: 0.0001      evaluation reward: 1.6511627906976745
episode: 86   score: 0.0   memory length: 16280   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.632183908045977
episode: 87   score: 2.0   memory length: 16478   epsilon: 1.0   steps: 1
98   lr: 0.0001      evaluation reward: 1.6363636363636365
episode: 88   score: 0.0   memory length: 16600   epsilon: 1.0   steps: 1
22   lr: 0.0001      evaluation reward: 1.6179775280898876
episode: 89   score: 2.0   memory length: 16800   epsilon: 1.0   steps: 2
00   lr: 0.0001      evaluation reward: 1.6222222222222222
episode: 90   score: 1.0   memory length: 16971   epsilon: 1.0   steps: 1
71   lr: 0.0001      evaluation reward: 1.6153846153846154
episode: 91   score: 1.0   memory length: 17141   epsilon: 1.0   steps: 1
70   lr: 0.0001      evaluation reward: 1.608695652173913
episode: 92   score: 3.0   memory length: 17390   epsilon: 1.0   steps: 2
49   lr: 0.0001      evaluation reward: 1.6236559139784945
episode: 93   score: 1.0   memory length: 17541   epsilon: 1.0   steps: 1
51   lr: 0.0001      evaluation reward: 1.6170212765957446
episode: 94   score: 3.0   memory length: 17767   epsilon: 1.0   steps: 2
26   lr: 0.0001      evaluation reward: 1.631578947368421
episode: 95   score: 3.0   memory length: 18019   epsilon: 1.0   steps: 2
52   lr: 0.0001      evaluation reward: 1.6458333333333333
episode: 96   score: 3.0   memory length: 18265   epsilon: 1.0   steps: 2
46   lr: 0.0001      evaluation reward: 1.6597938144329898
episode: 97   score: 3.0   memory length: 18480   epsilon: 1.0   steps: 2
15   lr: 0.0001      evaluation reward: 1.6734693877551021
episode: 98   score: 0.0   memory length: 18603   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.6565656565656566
episode: 99   score: 0.0   memory length: 18726   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.64
episode: 100  score: 0.0   memory length: 18849   epsilon: 1.0   steps:
123   lr: 0.0001      evaluation reward: 1.64
```

episode: 101	score: 0.0	memory length: 18971	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.62		
episode: 102	score: 0.0	memory length: 19093	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.59		
episode: 103	score: 0.0	memory length: 19215	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.53		
episode: 104	score: 4.0	memory length: 19491	epsilon: 1.0	steps:
276	lr: 0.0001	evaluation reward: 1.57		
episode: 105	score: 12.0	memory length: 20001	epsilon: 1.0	steps:
510	lr: 0.0001	evaluation reward: 1.68		
episode: 106	score: 2.0	memory length: 20200	epsilon: 1.0	steps:
199	lr: 0.0001	evaluation reward: 1.7		
episode: 107	score: 2.0	memory length: 20419	epsilon: 1.0	steps:
219	lr: 0.0001	evaluation reward: 1.71		
episode: 108	score: 2.0	memory length: 20636	epsilon: 1.0	steps:
217	lr: 0.0001	evaluation reward: 1.72		
episode: 109	score: 1.0	memory length: 20805	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.7		
episode: 110	score: 0.0	memory length: 20927	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.68		
episode: 111	score: 3.0	memory length: 21175	epsilon: 1.0	steps:
248	lr: 0.0001	evaluation reward: 1.68		
episode: 112	score: 2.0	memory length: 21373	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.7		
episode: 113	score: 2.0	memory length: 21591	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.71		
episode: 114	score: 4.0	memory length: 21907	epsilon: 1.0	steps:
316	lr: 0.0001	evaluation reward: 1.73		
episode: 115	score: 0.0	memory length: 22029	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.71		
episode: 116	score: 1.0	memory length: 22199	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.72		
episode: 117	score: 2.0	memory length: 22397	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.73		
episode: 118	score: 1.0	memory length: 22565	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.74		
episode: 119	score: 1.0	memory length: 22736	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.73		
episode: 120	score: 1.0	memory length: 22904	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.71		
episode: 121	score: 0.0	memory length: 23027	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.71		
episode: 122	score: 5.0	memory length: 23344	epsilon: 1.0	steps:
317	lr: 0.0001	evaluation reward: 1.76		
episode: 123	score: 4.0	memory length: 23661	epsilon: 1.0	steps:
317	lr: 0.0001	evaluation reward: 1.79		
episode: 124	score: 3.0	memory length: 23926	epsilon: 1.0	steps:
265	lr: 0.0001	evaluation reward: 1.79		
episode: 125	score: 3.0	memory length: 24172	epsilon: 1.0	steps:
246	lr: 0.0001	evaluation reward: 1.81		
episode: 126	score: 2.0	memory length: 24389	epsilon: 1.0	steps:
217	lr: 0.0001	evaluation reward: 1.8		
episode: 127	score: 2.0	memory length: 24587	epsilon: 1.0	steps:

198	lr: 0.0001	evaluation reward: 1.82		
episode: 128	score: 3.0	memory length: 24812	epsilon: 1.0	steps:
225	lr: 0.0001	evaluation reward: 1.84		
episode: 129	score: 2.0	memory length: 25009	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.84		
episode: 130	score: 1.0	memory length: 25178	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.83		
episode: 131	score: 2.0	memory length: 25375	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.85		
episode: 132	score: 3.0	memory length: 25622	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.88		
episode: 133	score: 1.0	memory length: 25791	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.89		
episode: 134	score: 3.0	memory length: 26039	epsilon: 1.0	steps:
248	lr: 0.0001	evaluation reward: 1.92		
episode: 135	score: 0.0	memory length: 26162	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.9		
episode: 136	score: 1.0	memory length: 26330	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.91		
episode: 137	score: 5.0	memory length: 26623	epsilon: 1.0	steps:
293	lr: 0.0001	evaluation reward: 1.94		
episode: 138	score: 1.0	memory length: 26792	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.93		
episode: 139	score: 1.0	memory length: 26942	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.91		
episode: 140	score: 0.0	memory length: 27064	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.89		
episode: 141	score: 1.0	memory length: 27214	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.87		
episode: 142	score: 4.0	memory length: 27488	epsilon: 1.0	steps:
274	lr: 0.0001	evaluation reward: 1.9		
episode: 143	score: 0.0	memory length: 27611	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.87		
episode: 144	score: 1.0	memory length: 27762	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.87		
episode: 145	score: 1.0	memory length: 27933	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.84		
episode: 146	score: 3.0	memory length: 28185	epsilon: 1.0	steps:
252	lr: 0.0001	evaluation reward: 1.87		
episode: 147	score: 3.0	memory length: 28411	epsilon: 1.0	steps:
226	lr: 0.0001	evaluation reward: 1.87		
episode: 148	score: 2.0	memory length: 28596	epsilon: 1.0	steps:
185	lr: 0.0001	evaluation reward: 1.87		
episode: 149	score: 0.0	memory length: 28719	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.85		
episode: 150	score: 1.0	memory length: 28870	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.85		
episode: 151	score: 2.0	memory length: 29068	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.85		
episode: 152	score: 3.0	memory length: 29317	epsilon: 1.0	steps:
249	lr: 0.0001	evaluation reward: 1.87		
episode: 153	score: 2.0	memory length: 29515	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.86		

episode: 154	score: 2.0	memory length: 29713	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.88		
episode: 155	score: 2.0	memory length: 29897	epsilon: 1.0	steps:
184	lr: 0.0001	evaluation reward: 1.88		
episode: 156	score: 0.0	memory length: 30019	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.85		
episode: 157	score: 3.0	memory length: 30266	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.85		
episode: 158	score: 0.0	memory length: 30389	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.78		
episode: 159	score: 0.0	memory length: 30511	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.76		
episode: 160	score: 2.0	memory length: 30729	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.77		
episode: 161	score: 3.0	memory length: 30975	epsilon: 1.0	steps:
246	lr: 0.0001	evaluation reward: 1.79		
episode: 162	score: 3.0	memory length: 31222	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.82		
episode: 163	score: 3.0	memory length: 31467	epsilon: 1.0	steps:
245	lr: 0.0001	evaluation reward: 1.82		
episode: 164	score: 1.0	memory length: 31618	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.81		
episode: 165	score: 3.0	memory length: 31863	epsilon: 1.0	steps:
245	lr: 0.0001	evaluation reward: 1.83		
episode: 166	score: 2.0	memory length: 32083	epsilon: 1.0	steps:
220	lr: 0.0001	evaluation reward: 1.84		
episode: 167	score: 3.0	memory length: 32330	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.85		
episode: 168	score: 2.0	memory length: 32531	epsilon: 1.0	steps:
201	lr: 0.0001	evaluation reward: 1.81		
episode: 169	score: 4.0	memory length: 32808	epsilon: 1.0	steps:
277	lr: 0.0001	evaluation reward: 1.85		
episode: 170	score: 2.0	memory length: 33026	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.87		
episode: 171	score: 1.0	memory length: 33177	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.87		
episode: 172	score: 1.0	memory length: 33347	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.86		
episode: 173	score: 4.0	memory length: 33641	epsilon: 1.0	steps:
294	lr: 0.0001	evaluation reward: 1.86		
episode: 174	score: 0.0	memory length: 33764	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.84		
episode: 175	score: 2.0	memory length: 33961	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.84		
episode: 176	score: 3.0	memory length: 34225	epsilon: 1.0	steps:
264	lr: 0.0001	evaluation reward: 1.87		
episode: 177	score: 3.0	memory length: 34436	epsilon: 1.0	steps:
211	lr: 0.0001	evaluation reward: 1.89		
episode: 178	score: 1.0	memory length: 34586	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.89		
episode: 179	score: 3.0	memory length: 34829	epsilon: 1.0	steps:
243	lr: 0.0001	evaluation reward: 1.91		
episode: 180	score: 3.0	memory length: 35057	epsilon: 1.0	steps:

228	lr: 0.0001	evaluation reward: 1.94		
episode: 181	score: 0.0	memory length: 35180	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.91		
episode: 182	score: 0.0	memory length: 35303	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.9		
episode: 183	score: 3.0	memory length: 35530	epsilon: 1.0	steps:
227	lr: 0.0001	evaluation reward: 1.91		
episode: 184	score: 1.0	memory length: 35699	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.9		
episode: 185	score: 5.0	memory length: 36025	epsilon: 1.0	steps:
326	lr: 0.0001	evaluation reward: 1.92		
episode: 186	score: 1.0	memory length: 36176	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.93		
episode: 187	score: 0.0	memory length: 36298	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.91		
episode: 188	score: 0.0	memory length: 36421	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.91		
episode: 189	score: 3.0	memory length: 36647	epsilon: 1.0	steps:
226	lr: 0.0001	evaluation reward: 1.92		
episode: 190	score: 2.0	memory length: 36845	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.93		
episode: 191	score: 1.0	memory length: 37013	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.93		
episode: 192	score: 1.0	memory length: 37184	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.91		
episode: 193	score: 0.0	memory length: 37306	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.9		
episode: 194	score: 2.0	memory length: 37525	epsilon: 1.0	steps:
219	lr: 0.0001	evaluation reward: 1.89		
episode: 195	score: 2.0	memory length: 37743	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.88		
episode: 196	score: 1.0	memory length: 37911	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.86		
episode: 197	score: 1.0	memory length: 38062	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.84		
episode: 198	score: 0.0	memory length: 38185	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.84		
episode: 199	score: 2.0	memory length: 38382	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.86		
episode: 200	score: 2.0	memory length: 38600	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.88		
episode: 201	score: 0.0	memory length: 38723	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.88		
episode: 202	score: 2.0	memory length: 38923	epsilon: 1.0	steps:
200	lr: 0.0001	evaluation reward: 1.9		
episode: 203	score: 1.0	memory length: 39075	epsilon: 1.0	steps:
152	lr: 0.0001	evaluation reward: 1.91		
episode: 204	score: 0.0	memory length: 39198	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.87		
episode: 205	score: 2.0	memory length: 39416	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.77		
episode: 206	score: 4.0	memory length: 39712	epsilon: 1.0	steps:
296	lr: 0.0001	evaluation reward: 1.79		

episode: 207	score: 0.0	memory length: 39834	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.77		
episode: 208	score: 2.0	memory length: 40049	epsilon: 1.0	steps:
215	lr: 0.0001	evaluation reward: 1.77		
episode: 209	score: 1.0	memory length: 40218	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.77		
episode: 210	score: 0.0	memory length: 40341	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.77		
episode: 211	score: 0.0	memory length: 40463	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.74		
episode: 212	score: 2.0	memory length: 40663	epsilon: 1.0	steps:
200	lr: 0.0001	evaluation reward: 1.74		
episode: 213	score: 4.0	memory length: 40980	epsilon: 1.0	steps:
317	lr: 0.0001	evaluation reward: 1.76		
episode: 214	score: 1.0	memory length: 41131	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.73		
episode: 215	score: 0.0	memory length: 41254	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.73		
episode: 216	score: 1.0	memory length: 41406	epsilon: 1.0	steps:
152	lr: 0.0001	evaluation reward: 1.73		
episode: 217	score: 0.0	memory length: 41529	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.71		
episode: 218	score: 0.0	memory length: 41651	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.7		
episode: 219	score: 0.0	memory length: 41773	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.69		
episode: 220	score: 3.0	memory length: 42016	epsilon: 1.0	steps:
243	lr: 0.0001	evaluation reward: 1.71		
episode: 221	score: 0.0	memory length: 42139	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.71		
episode: 222	score: 2.0	memory length: 42357	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.68		
episode: 223	score: 1.0	memory length: 42526	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.65		
episode: 224	score: 1.0	memory length: 42696	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.63		
episode: 225	score: 6.0	memory length: 43082	epsilon: 1.0	steps:
386	lr: 0.0001	evaluation reward: 1.66		
episode: 226	score: 0.0	memory length: 43204	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.64		
episode: 227	score: 0.0	memory length: 43326	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.62		
episode: 228	score: 2.0	memory length: 43524	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.61		
episode: 229	score: 0.0	memory length: 43646	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.59		
episode: 230	score: 2.0	memory length: 43861	epsilon: 1.0	steps:
215	lr: 0.0001	evaluation reward: 1.6		
episode: 231	score: 1.0	memory length: 44011	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.59		
episode: 232	score: 1.0	memory length: 44180	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.57		
episode: 233	score: 4.0	memory length: 44457	epsilon: 1.0	steps:

277	lr: 0.0001	evaluation reward: 1.6		
episode: 234	score: 1.0	memory length: 44628	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.58		
episode: 235	score: 2.0	memory length: 44844	epsilon: 1.0	steps:
216	lr: 0.0001	evaluation reward: 1.6		
episode: 236	score: 2.0	memory length: 45059	epsilon: 1.0	steps:
215	lr: 0.0001	evaluation reward: 1.61		
episode: 237	score: 2.0	memory length: 45277	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.58		
episode: 238	score: 0.0	memory length: 45399	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.57		
episode: 239	score: 3.0	memory length: 45646	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.59		
episode: 240	score: 7.0	memory length: 46060	epsilon: 1.0	steps:
414	lr: 0.0001	evaluation reward: 1.66		
episode: 241	score: 2.0	memory length: 46278	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.67		
episode: 242	score: 2.0	memory length: 46476	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.65		
episode: 243	score: 3.0	memory length: 46741	epsilon: 1.0	steps:
265	lr: 0.0001	evaluation reward: 1.68		
episode: 244	score: 2.0	memory length: 46942	epsilon: 1.0	steps:
201	lr: 0.0001	evaluation reward: 1.69		
episode: 245	score: 1.0	memory length: 47114	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.69		
episode: 246	score: 4.0	memory length: 47405	epsilon: 1.0	steps:
291	lr: 0.0001	evaluation reward: 1.7		
episode: 247	score: 1.0	memory length: 47573	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.68		
episode: 248	score: 0.0	memory length: 47696	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.66		
episode: 249	score: 3.0	memory length: 47943	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.69		
episode: 250	score: 0.0	memory length: 48065	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.68		
episode: 251	score: 1.0	memory length: 48236	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.67		
episode: 252	score: 1.0	memory length: 48406	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.65		
episode: 253	score: 1.0	memory length: 48575	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.64		
episode: 254	score: 1.0	memory length: 48745	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.63		
episode: 255	score: 1.0	memory length: 48917	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.62		
episode: 256	score: 1.0	memory length: 49085	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.63		
episode: 257	score: 1.0	memory length: 49236	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.61		
episode: 258	score: 3.0	memory length: 49462	epsilon: 1.0	steps:
226	lr: 0.0001	evaluation reward: 1.64		
episode: 259	score: 1.0	memory length: 49632	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.65		

episode: 260	score: 1.0	memory length: 49801	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.64		
episode: 261	score: 0.0	memory length: 49923	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.61		
episode: 262	score: 0.0	memory length: 50045	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.58		
episode: 263	score: 3.0	memory length: 50289	epsilon: 1.0	steps:
244	lr: 0.0001	evaluation reward: 1.58		
episode: 264	score: 2.0	memory length: 50486	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.59		
episode: 265	score: 2.0	memory length: 50705	epsilon: 1.0	steps:
219	lr: 0.0001	evaluation reward: 1.58		
episode: 266	score: 1.0	memory length: 50876	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.57		
episode: 267	score: 0.0	memory length: 50999	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.54		
episode: 268	score: 1.0	memory length: 51149	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.53		
episode: 269	score: 1.0	memory length: 51300	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.5		
episode: 270	score: 4.0	memory length: 51576	epsilon: 1.0	steps:
276	lr: 0.0001	evaluation reward: 1.52		
episode: 271	score: 0.0	memory length: 51699	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.51		
episode: 272	score: 2.0	memory length: 51918	epsilon: 1.0	steps:
219	lr: 0.0001	evaluation reward: 1.52		
episode: 273	score: 2.0	memory length: 52116	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.5		
episode: 274	score: 2.0	memory length: 52334	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.52		
episode: 275	score: 0.0	memory length: 52456	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.5		
episode: 276	score: 2.0	memory length: 52674	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.49		
episode: 277	score: 0.0	memory length: 52797	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.46		
episode: 278	score: 1.0	memory length: 52968	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.46		
episode: 279	score: 2.0	memory length: 53166	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.45		
episode: 280	score: 4.0	memory length: 53459	epsilon: 1.0	steps:
293	lr: 0.0001	evaluation reward: 1.46		
episode: 281	score: 1.0	memory length: 53610	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.47		
episode: 282	score: 0.0	memory length: 53732	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.47		
episode: 283	score: 0.0	memory length: 53854	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.44		
episode: 284	score: 4.0	memory length: 54109	epsilon: 1.0	steps:
255	lr: 0.0001	evaluation reward: 1.47		
episode: 285	score: 2.0	memory length: 54306	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.44		
episode: 286	score: 2.0	memory length: 54504	epsilon: 1.0	steps:

198	lr: 0.0001	evaluation reward: 1.45		
episode: 287	score: 2.0	memory length: 54702	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.47		
episode: 288	score: 3.0	memory length: 54948	epsilon: 1.0	steps:
246	lr: 0.0001	evaluation reward: 1.5		
episode: 289	score: 0.0	memory length: 55071	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.47		
episode: 290	score: 1.0	memory length: 55240	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.46		
episode: 291	score: 1.0	memory length: 55409	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.46		
episode: 292	score: 4.0	memory length: 55705	epsilon: 1.0	steps:
296	lr: 0.0001	evaluation reward: 1.49		
episode: 293	score: 3.0	memory length: 55968	epsilon: 1.0	steps:
263	lr: 0.0001	evaluation reward: 1.52		
episode: 294	score: 2.0	memory length: 56187	epsilon: 1.0	steps:
219	lr: 0.0001	evaluation reward: 1.52		
episode: 295	score: 2.0	memory length: 56385	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.52		
episode: 296	score: 3.0	memory length: 56632	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.54		
episode: 297	score: 2.0	memory length: 56850	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.55		
episode: 298	score: 5.0	memory length: 57176	epsilon: 1.0	steps:
326	lr: 0.0001	evaluation reward: 1.6		
episode: 299	score: 3.0	memory length: 57421	epsilon: 1.0	steps:
245	lr: 0.0001	evaluation reward: 1.61		
episode: 300	score: 0.0	memory length: 57543	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.59		
episode: 301	score: 1.0	memory length: 57712	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.6		
episode: 302	score: 1.0	memory length: 57881	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.59		
episode: 303	score: 0.0	memory length: 58003	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.58		
episode: 304	score: 1.0	memory length: 58154	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.59		
episode: 305	score: 2.0	memory length: 58352	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.59		
episode: 306	score: 1.0	memory length: 58505	epsilon: 1.0	steps:
153	lr: 0.0001	evaluation reward: 1.56		
episode: 307	score: 1.0	memory length: 58656	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.57		
episode: 308	score: 2.0	memory length: 58874	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.57		
episode: 309	score: 0.0	memory length: 58997	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.56		
episode: 310	score: 0.0	memory length: 59120	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.56		
episode: 311	score: 3.0	memory length: 59367	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.59		
episode: 312	score: 0.0	memory length: 59490	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.57		

episode: 313	score: 1.0	memory length: 59640	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.54		
episode: 314	score: 4.0	memory length: 59937	epsilon: 1.0	steps:
297	lr: 0.0001	evaluation reward: 1.57		
episode: 315	score: 1.0	memory length: 60087	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.58		
episode: 316	score: 0.0	memory length: 60210	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.57		
episode: 317	score: 0.0	memory length: 60333	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.57		
episode: 318	score: 3.0	memory length: 60580	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.6		
episode: 319	score: 5.0	memory length: 60892	epsilon: 1.0	steps:
312	lr: 0.0001	evaluation reward: 1.65		
episode: 320	score: 3.0	memory length: 61137	epsilon: 1.0	steps:
245	lr: 0.0001	evaluation reward: 1.65		
episode: 321	score: 2.0	memory length: 61335	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.67		
episode: 322	score: 2.0	memory length: 61533	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.67		
episode: 323	score: 3.0	memory length: 61778	epsilon: 1.0	steps:
245	lr: 0.0001	evaluation reward: 1.69		
episode: 324	score: 1.0	memory length: 61929	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.69		
episode: 325	score: 2.0	memory length: 62126	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.65		
episode: 326	score: 0.0	memory length: 62249	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.65		
episode: 327	score: 1.0	memory length: 62421	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.66		
episode: 328	score: 3.0	memory length: 62649	epsilon: 1.0	steps:
228	lr: 0.0001	evaluation reward: 1.67		
episode: 329	score: 2.0	memory length: 62847	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.69		
episode: 330	score: 0.0	memory length: 62969	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.67		
episode: 331	score: 2.0	memory length: 63150	epsilon: 1.0	steps:
181	lr: 0.0001	evaluation reward: 1.68		
episode: 332	score: 0.0	memory length: 63273	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.67		
episode: 333	score: 1.0	memory length: 63442	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.64		
episode: 334	score: 3.0	memory length: 63689	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.66		
episode: 335	score: 6.0	memory length: 64063	epsilon: 1.0	steps:
374	lr: 0.0001	evaluation reward: 1.7		
episode: 336	score: 1.0	memory length: 64213	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.69		
episode: 337	score: 0.0	memory length: 64335	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.67		
episode: 338	score: 1.0	memory length: 64503	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.68		
episode: 339	score: 2.0	memory length: 64700	epsilon: 1.0	steps:

197	lr: 0.0001	evaluation reward: 1.67		
episode: 340	score: 0.0	memory length: 64823	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.6		
episode: 341	score: 0.0	memory length: 64946	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.58		
episode: 342	score: 3.0	memory length: 65215	epsilon: 1.0	steps:
269	lr: 0.0001	evaluation reward: 1.59		
episode: 343	score: 2.0	memory length: 65416	epsilon: 1.0	steps:
201	lr: 0.0001	evaluation reward: 1.58		
episode: 344	score: 0.0	memory length: 65539	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.56		
episode: 345	score: 2.0	memory length: 65737	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.57		
episode: 346	score: 1.0	memory length: 65906	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.54		
episode: 347	score: 1.0	memory length: 66056	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.54		
episode: 348	score: 0.0	memory length: 66179	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.54		
episode: 349	score: 2.0	memory length: 66381	epsilon: 1.0	steps:
202	lr: 0.0001	evaluation reward: 1.53		
episode: 350	score: 1.0	memory length: 66550	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.54		
episode: 351	score: 0.0	memory length: 66672	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.53		
episode: 352	score: 2.0	memory length: 66872	epsilon: 1.0	steps:
200	lr: 0.0001	evaluation reward: 1.54		
episode: 353	score: 1.0	memory length: 67022	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.54		
episode: 354	score: 0.0	memory length: 67145	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.53		
episode: 355	score: 1.0	memory length: 67313	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.53		
episode: 356	score: 0.0	memory length: 67436	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.52		
episode: 357	score: 0.0	memory length: 67559	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.51		
episode: 358	score: 0.0	memory length: 67682	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.48		
episode: 359	score: 0.0	memory length: 67804	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.47		
episode: 360	score: 0.0	memory length: 67926	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.46		
episode: 361	score: 3.0	memory length: 68174	epsilon: 1.0	steps:
248	lr: 0.0001	evaluation reward: 1.49		
episode: 362	score: 0.0	memory length: 68297	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.49		
episode: 363	score: 1.0	memory length: 68448	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.47		
episode: 364	score: 3.0	memory length: 68674	epsilon: 1.0	steps:
226	lr: 0.0001	evaluation reward: 1.48		
episode: 365	score: 0.0	memory length: 68797	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.46		

episode: 366	score: 2.0	memory length: 69013	epsilon: 1.0	steps:
216	lr: 0.0001	evaluation reward: 1.47		
episode: 367	score: 7.0	memory length: 69431	epsilon: 1.0	steps:
418	lr: 0.0001	evaluation reward: 1.54		
episode: 368	score: 0.0	memory length: 69553	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.53		
episode: 369	score: 0.0	memory length: 69676	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.52		
episode: 370	score: 0.0	memory length: 69799	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.48		
episode: 371	score: 0.0	memory length: 69922	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.48		
episode: 372	score: 3.0	memory length: 70188	epsilon: 1.0	steps:
266	lr: 0.0001	evaluation reward: 1.49		
episode: 373	score: 3.0	memory length: 70436	epsilon: 1.0	steps:
248	lr: 0.0001	evaluation reward: 1.5		
episode: 374	score: 0.0	memory length: 70558	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.48		
episode: 375	score: 1.0	memory length: 70708	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.49		
episode: 376	score: 2.0	memory length: 70905	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.49		
episode: 377	score: 1.0	memory length: 71056	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.5		
episode: 378	score: 0.0	memory length: 71179	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.49		
episode: 379	score: 2.0	memory length: 71402	epsilon: 1.0	steps:
223	lr: 0.0001	evaluation reward: 1.49		
episode: 380	score: 1.0	memory length: 71571	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.46		
episode: 381	score: 2.0	memory length: 71789	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.47		
episode: 382	score: 6.0	memory length: 72162	epsilon: 1.0	steps:
373	lr: 0.0001	evaluation reward: 1.53		
episode: 383	score: 2.0	memory length: 72360	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.55		
episode: 384	score: 0.0	memory length: 72483	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.51		
episode: 385	score: 4.0	memory length: 72778	epsilon: 1.0	steps:
295	lr: 0.0001	evaluation reward: 1.53		
episode: 386	score: 3.0	memory length: 73048	epsilon: 1.0	steps:
270	lr: 0.0001	evaluation reward: 1.54		
episode: 387	score: 4.0	memory length: 73365	epsilon: 1.0	steps:
317	lr: 0.0001	evaluation reward: 1.56		
episode: 388	score: 4.0	memory length: 73641	epsilon: 1.0	steps:
276	lr: 0.0001	evaluation reward: 1.57		
episode: 389	score: 3.0	memory length: 73872	epsilon: 1.0	steps:
231	lr: 0.0001	evaluation reward: 1.6		
episode: 390	score: 0.0	memory length: 73995	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.59		
episode: 391	score: 1.0	memory length: 74165	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.59		
episode: 392	score: 1.0	memory length: 74334	epsilon: 1.0	steps:

169	lr: 0.0001	evaluation reward: 1.56		
episode: 393	score: 2.0	memory length: 74554	epsilon: 1.0	steps:
220	lr: 0.0001	evaluation reward: 1.55		
episode: 394	score: 0.0	memory length: 74677	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.53		
episode: 395	score: 3.0	memory length: 74905	epsilon: 1.0	steps:
228	lr: 0.0001	evaluation reward: 1.54		
episode: 396	score: 3.0	memory length: 75172	epsilon: 1.0	steps:
267	lr: 0.0001	evaluation reward: 1.54		
episode: 397	score: 2.0	memory length: 75370	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.54		
episode: 398	score: 0.0	memory length: 75492	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.49		
episode: 399	score: 0.0	memory length: 75615	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.46		
episode: 400	score: 1.0	memory length: 75783	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.47		
episode: 401	score: 1.0	memory length: 75936	epsilon: 1.0	steps:
153	lr: 0.0001	evaluation reward: 1.47		
episode: 402	score: 1.0	memory length: 76104	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.47		
episode: 403	score: 1.0	memory length: 76273	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.48		
episode: 404	score: 4.0	memory length: 76566	epsilon: 1.0	steps:
293	lr: 0.0001	evaluation reward: 1.51		
episode: 405	score: 2.0	memory length: 76783	epsilon: 1.0	steps:
217	lr: 0.0001	evaluation reward: 1.51		
episode: 406	score: 0.0	memory length: 76905	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.5		
episode: 407	score: 0.0	memory length: 77028	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.49		
episode: 408	score: 3.0	memory length: 77253	epsilon: 1.0	steps:
225	lr: 0.0001	evaluation reward: 1.5		
episode: 409	score: 2.0	memory length: 77450	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.52		
episode: 410	score: 0.0	memory length: 77573	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.52		
episode: 411	score: 2.0	memory length: 77770	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.51		
episode: 412	score: 1.0	memory length: 77921	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.52		
episode: 413	score: 0.0	memory length: 78044	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.51		
episode: 414	score: 3.0	memory length: 78292	epsilon: 1.0	steps:
248	lr: 0.0001	evaluation reward: 1.5		
episode: 415	score: 3.0	memory length: 78538	epsilon: 1.0	steps:
246	lr: 0.0001	evaluation reward: 1.52		
episode: 416	score: 5.0	memory length: 78862	epsilon: 1.0	steps:
324	lr: 0.0001	evaluation reward: 1.57		
episode: 417	score: 0.0	memory length: 78984	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.57		
episode: 418	score: 2.0	memory length: 79181	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.56		

episode: 419	score: 2.0	memory length: 79398	epsilon: 1.0	steps:
217	lr: 0.0001	evaluation reward: 1.53		
episode: 420	score: 0.0	memory length: 79520	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.5		
episode: 421	score: 2.0	memory length: 79717	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.5		
episode: 422	score: 1.0	memory length: 79889	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.49		
episode: 423	score: 1.0	memory length: 80061	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.47		
episode: 424	score: 0.0	memory length: 80183	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.46		
episode: 425	score: 0.0	memory length: 80306	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.44		
episode: 426	score: 0.0	memory length: 80429	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.44		
episode: 427	score: 3.0	memory length: 80678	epsilon: 1.0	steps:
249	lr: 0.0001	evaluation reward: 1.46		
episode: 428	score: 1.0	memory length: 80828	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.44		
episode: 429	score: 0.0	memory length: 80951	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.42		
episode: 430	score: 0.0	memory length: 81074	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.42		
episode: 431	score: 2.0	memory length: 81289	epsilon: 1.0	steps:
215	lr: 0.0001	evaluation reward: 1.42		
episode: 432	score: 2.0	memory length: 81487	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.44		
episode: 433	score: 1.0	memory length: 81638	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.44		
episode: 434	score: 0.0	memory length: 81761	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.41		
episode: 435	score: 0.0	memory length: 81884	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.35		
episode: 436	score: 2.0	memory length: 82081	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.36		
episode: 437	score: 0.0	memory length: 82204	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.36		
episode: 438	score: 0.0	memory length: 82326	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.35		
episode: 439	score: 2.0	memory length: 82523	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.35		
episode: 440	score: 0.0	memory length: 82646	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.35		
episode: 441	score: 2.0	memory length: 82844	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.37		
episode: 442	score: 1.0	memory length: 82997	epsilon: 1.0	steps:
153	lr: 0.0001	evaluation reward: 1.35		
episode: 443	score: 0.0	memory length: 83119	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.33		
episode: 444	score: 0.0	memory length: 83242	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.33		
episode: 445	score: 2.0	memory length: 83439	epsilon: 1.0	steps:

197	lr: 0.0001	evaluation reward: 1.33		
episode: 446	score: 5.0	memory length: 83763	epsilon: 1.0	steps:
324	lr: 0.0001	evaluation reward: 1.37		
episode: 447	score: 2.0	memory length: 83943	epsilon: 1.0	steps:
180	lr: 0.0001	evaluation reward: 1.38		
episode: 448	score: 2.0	memory length: 84141	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.4		
episode: 449	score: 0.0	memory length: 84263	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.38		
episode: 450	score: 2.0	memory length: 84465	epsilon: 1.0	steps:
202	lr: 0.0001	evaluation reward: 1.39		
episode: 451	score: 1.0	memory length: 84637	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.4		
episode: 452	score: 2.0	memory length: 84818	epsilon: 1.0	steps:
181	lr: 0.0001	evaluation reward: 1.4		
episode: 453	score: 3.0	memory length: 85044	epsilon: 1.0	steps:
226	lr: 0.0001	evaluation reward: 1.42		
episode: 454	score: 0.0	memory length: 85167	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.42		
episode: 455	score: 0.0	memory length: 85290	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.41		
episode: 456	score: 0.0	memory length: 85413	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.41		
episode: 457	score: 2.0	memory length: 85610	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.43		
episode: 458	score: 0.0	memory length: 85733	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.43		
episode: 459	score: 0.0	memory length: 85855	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.43		
episode: 460	score: 2.0	memory length: 86073	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.45		
episode: 461	score: 3.0	memory length: 86299	epsilon: 1.0	steps:
226	lr: 0.0001	evaluation reward: 1.45		
episode: 462	score: 0.0	memory length: 86421	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.45		
episode: 463	score: 3.0	memory length: 86668	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.47		
episode: 464	score: 0.0	memory length: 86791	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.44		
episode: 465	score: 3.0	memory length: 87038	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.47		
episode: 466	score: 1.0	memory length: 87189	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.46		
episode: 467	score: 3.0	memory length: 87436	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.42		
episode: 468	score: 1.0	memory length: 87605	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.43		
episode: 469	score: 7.0	memory length: 87904	epsilon: 1.0	steps:
299	lr: 0.0001	evaluation reward: 1.5		
episode: 470	score: 0.0	memory length: 88026	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.5		
episode: 471	score: 1.0	memory length: 88197	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.51		

episode: 472	score: 1.0	memory length: 88366	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.49		
episode: 473	score: 0.0	memory length: 88488	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.46		
episode: 474	score: 0.0	memory length: 88610	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.46		
episode: 475	score: 2.0	memory length: 88811	epsilon: 1.0	steps:
201	lr: 0.0001	evaluation reward: 1.47		
episode: 476	score: 1.0	memory length: 88980	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.46		
episode: 477	score: 2.0	memory length: 89178	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.47		
episode: 478	score: 0.0	memory length: 89301	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.47		
episode: 479	score: 4.0	memory length: 89577	epsilon: 1.0	steps:
276	lr: 0.0001	evaluation reward: 1.49		
episode: 480	score: 3.0	memory length: 89823	epsilon: 1.0	steps:
246	lr: 0.0001	evaluation reward: 1.51		
episode: 481	score: 2.0	memory length: 90007	epsilon: 1.0	steps:
184	lr: 0.0001	evaluation reward: 1.51		
episode: 482	score: 1.0	memory length: 90178	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.46		
episode: 483	score: 0.0	memory length: 90301	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.44		
episode: 484	score: 1.0	memory length: 90470	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.45		
episode: 485	score: 2.0	memory length: 90667	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.43		
episode: 486	score: 1.0	memory length: 90839	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.41		
episode: 487	score: 1.0	memory length: 90990	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.38		
episode: 488	score: 0.0	memory length: 91113	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.34		
episode: 489	score: 0.0	memory length: 91236	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.31		
episode: 490	score: 2.0	memory length: 91454	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.33		
episode: 491	score: 0.0	memory length: 91577	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.32		
episode: 492	score: 0.0	memory length: 91700	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.31		
episode: 493	score: 5.0	memory length: 92011	epsilon: 1.0	steps:
311	lr: 0.0001	evaluation reward: 1.34		
episode: 494	score: 2.0	memory length: 92229	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.36		
episode: 495	score: 2.0	memory length: 92409	epsilon: 1.0	steps:
180	lr: 0.0001	evaluation reward: 1.35		
episode: 496	score: 2.0	memory length: 92629	epsilon: 1.0	steps:
220	lr: 0.0001	evaluation reward: 1.34		
episode: 497	score: 1.0	memory length: 92797	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.33		
episode: 498	score: 1.0	memory length: 92965	epsilon: 1.0	steps:

168	lr: 0.0001	evaluation reward: 1.34		
episode: 499	score: 2.0	memory length: 93185	epsilon: 1.0	steps:
220	lr: 0.0001	evaluation reward: 1.36		
episode: 500	score: 3.0	memory length: 93412	epsilon: 1.0	steps:
227	lr: 0.0001	evaluation reward: 1.38		
episode: 501	score: 2.0	memory length: 93630	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.39		
episode: 502	score: 2.0	memory length: 93828	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.4		
episode: 503	score: 2.0	memory length: 94026	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.41		
episode: 504	score: 0.0	memory length: 94149	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.37		
episode: 505	score: 0.0	memory length: 94271	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.35		
episode: 506	score: 0.0	memory length: 94394	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.35		
episode: 507	score: 3.0	memory length: 94640	epsilon: 1.0	steps:
246	lr: 0.0001	evaluation reward: 1.38		
episode: 508	score: 2.0	memory length: 94855	epsilon: 1.0	steps:
215	lr: 0.0001	evaluation reward: 1.37		
episode: 509	score: 4.0	memory length: 95147	epsilon: 1.0	steps:
292	lr: 0.0001	evaluation reward: 1.39		
episode: 510	score: 1.0	memory length: 95297	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.4		
episode: 511	score: 1.0	memory length: 95465	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.39		
episode: 512	score: 0.0	memory length: 95587	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.38		
episode: 513	score: 0.0	memory length: 95709	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.38		
episode: 514	score: 2.0	memory length: 95927	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.37		
episode: 515	score: 3.0	memory length: 96176	epsilon: 1.0	steps:
249	lr: 0.0001	evaluation reward: 1.37		
episode: 516	score: 1.0	memory length: 96344	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.33		
episode: 517	score: 0.0	memory length: 96467	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.33		
episode: 518	score: 2.0	memory length: 96667	epsilon: 1.0	steps:
200	lr: 0.0001	evaluation reward: 1.33		
episode: 519	score: 3.0	memory length: 96878	epsilon: 1.0	steps:
211	lr: 0.0001	evaluation reward: 1.34		
episode: 520	score: 0.0	memory length: 97001	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.34		
episode: 521	score: 1.0	memory length: 97171	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.33		
episode: 522	score: 3.0	memory length: 97398	epsilon: 1.0	steps:
227	lr: 0.0001	evaluation reward: 1.35		
episode: 523	score: 3.0	memory length: 97646	epsilon: 1.0	steps:
248	lr: 0.0001	evaluation reward: 1.37		
episode: 524	score: 1.0	memory length: 97815	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.38		

```
episode: 525  score: 2.0  memory length: 98015  epsilon: 1.0  steps:
200  lr: 0.0001  evaluation reward: 1.4
episode: 526  score: 6.0  memory length: 98388  epsilon: 1.0  steps:
373  lr: 0.0001  evaluation reward: 1.46
episode: 527  score: 2.0  memory length: 98586  epsilon: 1.0  steps:
198  lr: 0.0001  evaluation reward: 1.45
episode: 528  score: 0.0  memory length: 98709  epsilon: 1.0  steps:
123  lr: 0.0001  evaluation reward: 1.44
episode: 529  score: 2.0  memory length: 98906  epsilon: 1.0  steps:
197  lr: 0.0001  evaluation reward: 1.46
episode: 530  score: 2.0  memory length: 99087  epsilon: 1.0  steps:
181  lr: 0.0001  evaluation reward: 1.48
episode: 531  score: 2.0  memory length: 99303  epsilon: 1.0  steps:
216  lr: 0.0001  evaluation reward: 1.48
episode: 532  score: 2.0  memory length: 99501  epsilon: 1.0  steps:
198  lr: 0.0001  evaluation reward: 1.48
episode: 533  score: 3.0  memory length: 99727  epsilon: 1.0  steps:
226  lr: 0.0001  evaluation reward: 1.5
episode: 534  score: 0.0  memory length: 99850  epsilon: 1.0  steps:
123  lr: 0.0001  evaluation reward: 1.5
episode: 535  score: 0.0  memory length: 99973  epsilon: 1.0  steps:
123  lr: 0.0001  evaluation reward: 1.5
```

```
/home/wxzhang/assignment5/memory.py:30: FutureWarning: The input object of
type 'Tensor' is an array-like implementing one of the corresponding proto
cols (`__array__`, `__array_interface__` or `__array_struct__`); but not a
sequence (or 0-D). In the future, this object will be coerced as if it was
first converted using `np.array(obj)`. To retain the old behaviour, you ha
ve to either modify the type 'Tensor', or assign to an empty array created
with `np.empty(correct_shape, dtype=object)`.
```

```
sample = np.array(sample, dtype=object)
```

```
/home/wxzhang/assignment5/agent.py:67: FutureWarning: The input object of
type 'Tensor' is an array-like implementing one of the corresponding proto
cols (`__array__`, `__array_interface__` or `__array_struct__`); but not a
sequence (or 0-D). In the future, this object will be coerced as if it was
first converted using `np.array(obj)`. To retain the old behaviour, you ha
ve to either modify the type 'Tensor', or assign to an empty array created
with `np.empty(correct_shape, dtype=object)`.
```

```
mini_batch = np.array(mini_batch, dtype=object).transpose()
```

```
episode: 536  score: 0.0  memory length: 100096  epsilon: 0.99980794000
00042  steps: 123  lr: 0.0001  evaluation reward: 1.48
episode: 537  score: 2.0  memory length: 100316  epsilon: 0.99937234000
00136  steps: 220  lr: 0.0001  evaluation reward: 1.5
episode: 538  score: 2.0  memory length: 100533  epsilon: 0.99894268000
0023  steps: 217  lr: 0.0001  evaluation reward: 1.52
episode: 539  score: 1.0  memory length: 100702  epsilon: 0.99860806000
00302  steps: 169  lr: 0.0001  evaluation reward: 1.51
episode: 540  score: 0.0  memory length: 100825  epsilon: 0.99836452000
00355  steps: 123  lr: 0.0001  evaluation reward: 1.51
episode: 541  score: 1.0  memory length: 100975  epsilon: 0.99806752000
0042  steps: 150  lr: 0.0001  evaluation reward: 1.5
episode: 542  score: 2.0  memory length: 101173  epsilon: 0.99767548000
00505  steps: 198  lr: 0.0001  evaluation reward: 1.51
episode: 543  score: 1.0  memory length: 101343  epsilon: 0.99733888000
```

```
00578  steps: 170  lr: 0.0001  evaluation reward: 1.52
episode: 544  score: 3.0  memory length: 101608  epsilon: 0.99681418000
00692  steps: 265  lr: 0.0001  evaluation reward: 1.55
episode: 545  score: 3.0  memory length: 101853  epsilon: 0.99632908000
00797  steps: 245  lr: 0.0001  evaluation reward: 1.56
episode: 546  score: 2.0  memory length: 102074  epsilon: 0.99589150000
00892  steps: 221  lr: 0.0001  evaluation reward: 1.53
episode: 547  score: 3.0  memory length: 102303  epsilon: 0.99543808000
0099  steps: 229  lr: 0.0001  evaluation reward: 1.54
episode: 548  score: 0.0  memory length: 102426  epsilon: 0.99519454000
01043  steps: 123  lr: 0.0001  evaluation reward: 1.52
episode: 549  score: 3.0  memory length: 102672  epsilon: 0.99470746000
01149  steps: 246  lr: 0.0001  evaluation reward: 1.55
episode: 550  score: 1.0  memory length: 102823  epsilon: 0.99440848000
01214  steps: 151  lr: 0.0001  evaluation reward: 1.54
episode: 551  score: 2.0  memory length: 103042  epsilon: 0.99397486000
01308  steps: 219  lr: 0.0001  evaluation reward: 1.55
episode: 552  score: 2.0  memory length: 103262  epsilon: 0.99353926000
01403  steps: 220  lr: 0.0001  evaluation reward: 1.55
episode: 553  score: 4.0  memory length: 103578  epsilon: 0.99291358000
01538  steps: 316  lr: 0.0001  evaluation reward: 1.56
episode: 554  score: 2.0  memory length: 103778  epsilon: 0.99251758000
01624  steps: 200  lr: 0.0001  evaluation reward: 1.58
episode: 555  score: 0.0  memory length: 103901  epsilon: 0.99227404000
01677  steps: 123  lr: 0.0001  evaluation reward: 1.58
episode: 556  score: 3.0  memory length: 104145  epsilon: 0.99179092000
01782  steps: 244  lr: 0.0001  evaluation reward: 1.61
episode: 557  score: 0.0  memory length: 104268  epsilon: 0.99154738000
01835  steps: 123  lr: 0.0001  evaluation reward: 1.59
episode: 558  score: 1.0  memory length: 104437  epsilon: 0.99121276000
01908  steps: 169  lr: 0.0001  evaluation reward: 1.6
episode: 559  score: 2.0  memory length: 104635  epsilon: 0.99082072000
01993  steps: 198  lr: 0.0001  evaluation reward: 1.62
episode: 560  score: 0.0  memory length: 104757  epsilon: 0.99057916000
02045  steps: 122  lr: 0.0001  evaluation reward: 1.6
episode: 561  score: 4.0  memory length: 105032  epsilon: 0.99003466000
02163  steps: 275  lr: 0.0001  evaluation reward: 1.61
episode: 562  score: 1.0  memory length: 105201  epsilon: 0.98970004000
02236  steps: 169  lr: 0.0001  evaluation reward: 1.62
episode: 563  score: 4.0  memory length: 105475  epsilon: 0.98915752000
02354  steps: 274  lr: 0.0001  evaluation reward: 1.63
episode: 564  score: 1.0  memory length: 105647  epsilon: 0.98881696000
02428  steps: 172  lr: 0.0001  evaluation reward: 1.64
episode: 565  score: 0.0  memory length: 105770  epsilon: 0.98857342000
02481  steps: 123  lr: 0.0001  evaluation reward: 1.61
episode: 566  score: 0.0  memory length: 105893  epsilon: 0.98832988000
02533  steps: 123  lr: 0.0001  evaluation reward: 1.6
episode: 567  score: 4.0  memory length: 106187  epsilon: 0.98774776000
0266  steps: 294  lr: 0.0001  evaluation reward: 1.61
episode: 568  score: 0.0  memory length: 106310  epsilon: 0.98750422000
02713  steps: 123  lr: 0.0001  evaluation reward: 1.6
episode: 569  score: 2.0  memory length: 106528  epsilon: 0.98707258000
02806  steps: 218  lr: 0.0001  evaluation reward: 1.55
```

```
episode: 570  score: 1.0  memory length: 106679  epsilon: 0.98677360000
02871  steps: 151  lr: 0.0001  evaluation reward: 1.56
episode: 571  score: 1.0  memory length: 106830  epsilon: 0.98647462000
02936  steps: 151  lr: 0.0001  evaluation reward: 1.56
episode: 572  score: 2.0  memory length: 107028  epsilon: 0.98608258000
03021  steps: 198  lr: 0.0001  evaluation reward: 1.57
episode: 573  score: 2.0  memory length: 107209  epsilon: 0.98572420000
03099  steps: 181  lr: 0.0001  evaluation reward: 1.59
episode: 574  score: 1.0  memory length: 107379  epsilon: 0.98538760000
03172  steps: 170  lr: 0.0001  evaluation reward: 1.6
episode: 575  score: 3.0  memory length: 107644  epsilon: 0.98486290000
03286  steps: 265  lr: 0.0001  evaluation reward: 1.61
episode: 576  score: 4.0  memory length: 107939  epsilon: 0.98427880000
03413  steps: 295  lr: 0.0001  evaluation reward: 1.64
episode: 577  score: 4.0  memory length: 108216  epsilon: 0.98373034000
03532  steps: 277  lr: 0.0001  evaluation reward: 1.66
episode: 578  score: 2.0  memory length: 108414  epsilon: 0.98333830000
03617  steps: 198  lr: 0.0001  evaluation reward: 1.68
episode: 579  score: 1.0  memory length: 108582  epsilon: 0.98300566000
03689  steps: 168  lr: 0.0001  evaluation reward: 1.65
episode: 580  score: 0.0  memory length: 108705  epsilon: 0.98276212000
03742  steps: 123  lr: 0.0001  evaluation reward: 1.62
episode: 581  score: 4.0  memory length: 108998  epsilon: 0.98218198000
03868  steps: 293  lr: 0.0001  evaluation reward: 1.64
episode: 582  score: 3.0  memory length: 109243  epsilon: 0.98169688000
03973  steps: 245  lr: 0.0001  evaluation reward: 1.66
episode: 583  score: 0.0  memory length: 109366  epsilon: 0.98145334000
04026  steps: 123  lr: 0.0001  evaluation reward: 1.66
episode: 584  score: 0.0  memory length: 109489  epsilon: 0.98120980000
04079  steps: 123  lr: 0.0001  evaluation reward: 1.65
episode: 585  score: 0.0  memory length: 109611  epsilon: 0.98096824000
04132  steps: 122  lr: 0.0001  evaluation reward: 1.63
episode: 586  score: 1.0  memory length: 109779  epsilon: 0.98063560000
04204  steps: 168  lr: 0.0001  evaluation reward: 1.63
episode: 587  score: 2.0  memory length: 109999  epsilon: 0.98020000000
04298  steps: 220  lr: 0.0001  evaluation reward: 1.64
episode: 588  score: 0.0  memory length: 110122  epsilon: 0.97995646000
04351  steps: 123  lr: 0.0001  evaluation reward: 1.64
episode: 589  score: 3.0  memory length: 110388  epsilon: 0.97942978000
04466  steps: 266  lr: 0.0001  evaluation reward: 1.67
episode: 590  score: 1.0  memory length: 110557  epsilon: 0.97909516000
04538  steps: 169  lr: 0.0001  evaluation reward: 1.66
episode: 591  score: 2.0  memory length: 110754  epsilon: 0.97870510000
04623  steps: 197  lr: 0.0001  evaluation reward: 1.68
episode: 592  score: 1.0  memory length: 110924  epsilon: 0.97836850000
04696  steps: 170  lr: 0.0001  evaluation reward: 1.69
episode: 593  score: 2.0  memory length: 111143  epsilon: 0.97793488000
0479  steps: 219  lr: 0.0001  evaluation reward: 1.66
episode: 594  score: 0.0  memory length: 111266  epsilon: 0.97769134000
04843  steps: 123  lr: 0.0001  evaluation reward: 1.64
episode: 595  score: 3.0  memory length: 111511  epsilon: 0.97720624000
04948  steps: 245  lr: 0.0001  evaluation reward: 1.65
episode: 596  score: 1.0  memory length: 111680  epsilon: 0.97687162000
```

```
05021  steps: 169    lr: 0.0001    evaluation reward: 1.64
episode: 597  score: 1.0  memory length: 111848  epsilon: 0.97653898000
05093  steps: 168    lr: 0.0001    evaluation reward: 1.64
episode: 598  score: 0.0  memory length: 111971  epsilon: 0.97629544000
05146  steps: 123    lr: 0.0001    evaluation reward: 1.63
episode: 599  score: 0.0  memory length: 112093  epsilon: 0.97605388000
05198  steps: 122    lr: 0.0001    evaluation reward: 1.61
episode: 600  score: 0.0  memory length: 112216  epsilon: 0.97581034000
05251  steps: 123    lr: 0.0001    evaluation reward: 1.58
episode: 601  score: 4.0  memory length: 112491  epsilon: 0.97526584000
0537   steps: 275    lr: 0.0001    evaluation reward: 1.6
episode: 602  score: 3.0  memory length: 112702  epsilon: 0.97484806000
0546   steps: 211    lr: 0.0001    evaluation reward: 1.61
episode: 603  score: 2.0  memory length: 112920  epsilon: 0.97441642000
05554  steps: 218    lr: 0.0001    evaluation reward: 1.61
episode: 604  score: 0.0  memory length: 113043  epsilon: 0.97417288000
05607  steps: 123    lr: 0.0001    evaluation reward: 1.61
episode: 605  score: 1.0  memory length: 113212  epsilon: 0.97383826000
0568   steps: 169    lr: 0.0001    evaluation reward: 1.62
episode: 606  score: 2.0  memory length: 113409  epsilon: 0.97344820000
05764  steps: 197    lr: 0.0001    evaluation reward: 1.64
episode: 607  score: 2.0  memory length: 113627  epsilon: 0.97301656000
05858  steps: 218    lr: 0.0001    evaluation reward: 1.63
episode: 608  score: 2.0  memory length: 113846  epsilon: 0.97258294000
05952  steps: 219    lr: 0.0001    evaluation reward: 1.63
episode: 609  score: 3.0  memory length: 114072  epsilon: 0.97213546000
06049  steps: 226    lr: 0.0001    evaluation reward: 1.62
episode: 610  score: 0.0  memory length: 114195  epsilon: 0.97189192000
06102  steps: 123    lr: 0.0001    evaluation reward: 1.61
episode: 611  score: 1.0  memory length: 114363  epsilon: 0.97155928000
06174  steps: 168    lr: 0.0001    evaluation reward: 1.61
episode: 612  score: 3.0  memory length: 114611  epsilon: 0.97106824000
06281  steps: 248    lr: 0.0001    evaluation reward: 1.64
episode: 613  score: 0.0  memory length: 114733  epsilon: 0.97082668000
06333  steps: 122    lr: 0.0001    evaluation reward: 1.64
episode: 614  score: 3.0  memory length: 114962  epsilon: 0.97037326000
06432  steps: 229    lr: 0.0001    evaluation reward: 1.65
episode: 615  score: 0.0  memory length: 115085  epsilon: 0.97012972000
06485  steps: 123    lr: 0.0001    evaluation reward: 1.62
episode: 616  score: 1.0  memory length: 115236  epsilon: 0.96983074000
0655   steps: 151    lr: 0.0001    evaluation reward: 1.62
episode: 617  score: 2.0  memory length: 115436  epsilon: 0.96943474000
06635  steps: 200    lr: 0.0001    evaluation reward: 1.64
episode: 618  score: 2.0  memory length: 115634  epsilon: 0.96904270000
0672   steps: 198    lr: 0.0001    evaluation reward: 1.64
episode: 619  score: 3.0  memory length: 115901  epsilon: 0.96851404000
06835  steps: 267    lr: 0.0001    evaluation reward: 1.64
episode: 620  score: 2.0  memory length: 116101  epsilon: 0.96811804000
06921  steps: 200    lr: 0.0001    evaluation reward: 1.66
episode: 621  score: 0.0  memory length: 116224  epsilon: 0.96787450000
06974  steps: 123    lr: 0.0001    evaluation reward: 1.65
episode: 622  score: 1.0  memory length: 116392  epsilon: 0.96754186000
07046  steps: 168    lr: 0.0001    evaluation reward: 1.63
```

```
episode: 623   score: 0.0   memory length: 116515   epsilon: 0.96729832000
07099   steps: 123   lr: 0.0001   evaluation reward: 1.6
episode: 624   score: 0.0   memory length: 116637   epsilon: 0.96705676000
07152   steps: 122   lr: 0.0001   evaluation reward: 1.59
episode: 625   score: 1.0   memory length: 116787   epsilon: 0.96675976000
07216   steps: 150   lr: 0.0001   evaluation reward: 1.58
episode: 626   score: 1.0   memory length: 116938   epsilon: 0.96646078000
07281   steps: 151   lr: 0.0001   evaluation reward: 1.53
episode: 627   score: 2.0   memory length: 117157   epsilon: 0.96602716000
07375   steps: 219   lr: 0.0001   evaluation reward: 1.53
episode: 628   score: 2.0   memory length: 117375   epsilon: 0.96559552000
07469   steps: 218   lr: 0.0001   evaluation reward: 1.55
episode: 629   score: 0.0   memory length: 117497   epsilon: 0.96535396000
07521   steps: 122   lr: 0.0001   evaluation reward: 1.53
episode: 630   score: 2.0   memory length: 117694   epsilon: 0.96496390000
07606   steps: 197   lr: 0.0001   evaluation reward: 1.53
episode: 631   score: 0.0   memory length: 117817   epsilon: 0.96472036000
07659   steps: 123   lr: 0.0001   evaluation reward: 1.51
episode: 632   score: 0.0   memory length: 117939   epsilon: 0.96447880000
07711   steps: 122   lr: 0.0001   evaluation reward: 1.49
episode: 633   score: 2.0   memory length: 118155   epsilon: 0.96405112000
07804   steps: 216   lr: 0.0001   evaluation reward: 1.48
episode: 634   score: 0.0   memory length: 118278   epsilon: 0.96380758000
07857   steps: 123   lr: 0.0001   evaluation reward: 1.48
episode: 635   score: 2.0   memory length: 118496   epsilon: 0.96337594000
07951   steps: 218   lr: 0.0001   evaluation reward: 1.5
episode: 636   score: 0.0   memory length: 118619   epsilon: 0.96313240000
08004   steps: 123   lr: 0.0001   evaluation reward: 1.5
episode: 637   score: 2.0   memory length: 118835   epsilon: 0.96270472000
08096   steps: 216   lr: 0.0001   evaluation reward: 1.5
episode: 638   score: 3.0   memory length: 119063   epsilon: 0.96225328000
08194   steps: 228   lr: 0.0001   evaluation reward: 1.51
episode: 639   score: 2.0   memory length: 119283   epsilon: 0.96181768000
08289   steps: 220   lr: 0.0001   evaluation reward: 1.52
episode: 640   score: 1.0   memory length: 119453   epsilon: 0.96148108000
08362   steps: 170   lr: 0.0001   evaluation reward: 1.53
episode: 641   score: 0.0   memory length: 119575   epsilon: 0.96123952000
08415   steps: 122   lr: 0.0001   evaluation reward: 1.52
episode: 642   score: 1.0   memory length: 119746   epsilon: 0.96090094000
08488   steps: 171   lr: 0.0001   evaluation reward: 1.51
episode: 643   score: 2.0   memory length: 119943   epsilon: 0.96051088000
08573   steps: 197   lr: 0.0001   evaluation reward: 1.52
episode: 644   score: 6.0   memory length: 120319   epsilon: 0.95976640000
08734   steps: 376   lr: 0.0001   evaluation reward: 1.55
episode: 645   score: 1.0   memory length: 120470   epsilon: 0.95946742000
08799   steps: 151   lr: 0.0001   evaluation reward: 1.53
episode: 646   score: 3.0   memory length: 120695   epsilon: 0.95902192000
08896   steps: 225   lr: 0.0001   evaluation reward: 1.54
episode: 647   score: 0.0   memory length: 120817   epsilon: 0.95878036000
08948   steps: 122   lr: 0.0001   evaluation reward: 1.51
episode: 648   score: 0.0   memory length: 120940   epsilon: 0.95853682000
09001   steps: 123   lr: 0.0001   evaluation reward: 1.51
episode: 649   score: 3.0   memory length: 121152   epsilon: 0.95811706000
```

```
09092  steps: 212  lr: 0.0001  evaluation reward: 1.51
episode: 650  score: 0.0  memory length: 121274  epsilon: 0.95787550000
09145  steps: 122  lr: 0.0001  evaluation reward: 1.5
episode: 651  score: 2.0  memory length: 121472  epsilon: 0.95748346000
0923  steps: 198  lr: 0.0001  evaluation reward: 1.5
episode: 652  score: 2.0  memory length: 121670  epsilon: 0.95709142000
09315  steps: 198  lr: 0.0001  evaluation reward: 1.5
episode: 653  score: 0.0  memory length: 121792  epsilon: 0.95684986000
09367  steps: 122  lr: 0.0001  evaluation reward: 1.46
episode: 654  score: 2.0  memory length: 122011  epsilon: 0.95641624000
09462  steps: 219  lr: 0.0001  evaluation reward: 1.46
episode: 655  score: 4.0  memory length: 122306  epsilon: 0.95583214000
09588  steps: 295  lr: 0.0001  evaluation reward: 1.5
episode: 656  score: 1.0  memory length: 122478  epsilon: 0.95549158000
09662  steps: 172  lr: 0.0001  evaluation reward: 1.48
episode: 657  score: 4.0  memory length: 122791  epsilon: 0.95487184000
09797  steps: 313  lr: 0.0001  evaluation reward: 1.52
episode: 658  score: 4.0  memory length: 123083  epsilon: 0.95429368000
09922  steps: 292  lr: 0.0001  evaluation reward: 1.55
episode: 659  score: 1.0  memory length: 123254  epsilon: 0.95395510000
09996  steps: 171  lr: 0.0001  evaluation reward: 1.54
episode: 660  score: 5.0  memory length: 123619  epsilon: 0.95323240000
10153  steps: 365  lr: 0.0001  evaluation reward: 1.59
episode: 661  score: 3.0  memory length: 123845  epsilon: 0.95278492000
1025  steps: 226  lr: 0.0001  evaluation reward: 1.58
episode: 662  score: 0.0  memory length: 123967  epsilon: 0.95254336000
10302  steps: 122  lr: 0.0001  evaluation reward: 1.57
episode: 663  score: 2.0  memory length: 124165  epsilon: 0.95215132000
10387  steps: 198  lr: 0.0001  evaluation reward: 1.55
episode: 664  score: 0.0  memory length: 124287  epsilon: 0.95190976000
1044  steps: 122  lr: 0.0001  evaluation reward: 1.54
episode: 665  score: 2.0  memory length: 124504  epsilon: 0.95148010000
10533  steps: 217  lr: 0.0001  evaluation reward: 1.56
episode: 666  score: 2.0  memory length: 124722  epsilon: 0.95104846000
10627  steps: 218  lr: 0.0001  evaluation reward: 1.58
episode: 667  score: 3.0  memory length: 124988  epsilon: 0.95052178000
10741  steps: 266  lr: 0.0001  evaluation reward: 1.57
episode: 668  score: 1.0  memory length: 125156  epsilon: 0.95018914000
10813  steps: 168  lr: 0.0001  evaluation reward: 1.58
episode: 669  score: 1.0  memory length: 125325  epsilon: 0.94985452000
10886  steps: 169  lr: 0.0001  evaluation reward: 1.57
episode: 670  score: 1.0  memory length: 125497  epsilon: 0.94951396000
1096  steps: 172  lr: 0.0001  evaluation reward: 1.57
episode: 671  score: 3.0  memory length: 125746  epsilon: 0.94902094000
11067  steps: 249  lr: 0.0001  evaluation reward: 1.59
episode: 672  score: 1.0  memory length: 125917  epsilon: 0.94868236000
11141  steps: 171  lr: 0.0001  evaluation reward: 1.58
episode: 673  score: 0.0  memory length: 126040  epsilon: 0.94843882000
11193  steps: 123  lr: 0.0001  evaluation reward: 1.56
episode: 674  score: 1.0  memory length: 126190  epsilon: 0.94814182000
11258  steps: 150  lr: 0.0001  evaluation reward: 1.56
episode: 675  score: 0.0  memory length: 126313  epsilon: 0.94789828000
11311  steps: 123  lr: 0.0001  evaluation reward: 1.53
```

```
episode: 676   score: 2.0   memory length: 126511   epsilon: 0.94750624000
11396   steps: 198   lr: 0.0001   evaluation reward: 1.51
episode: 677   score: 0.0   memory length: 126634   epsilon: 0.94726270000
11449   steps: 123   lr: 0.0001   evaluation reward: 1.47
episode: 678   score: 1.0   memory length: 126803   epsilon: 0.94692808000
11521   steps: 169   lr: 0.0001   evaluation reward: 1.46
episode: 679   score: 8.0   memory length: 127263   epsilon: 0.94601728000
11719   steps: 460   lr: 0.0001   evaluation reward: 1.53
episode: 680   score: 2.0   memory length: 127461   epsilon: 0.94562524000
11804   steps: 198   lr: 0.0001   evaluation reward: 1.55
episode: 681   score: 1.0   memory length: 127629   epsilon: 0.94529260000
11876   steps: 168   lr: 0.0001   evaluation reward: 1.52
episode: 682   score: 1.0   memory length: 127798   epsilon: 0.94495798000
11949   steps: 169   lr: 0.0001   evaluation reward: 1.5
episode: 683   score: 2.0   memory length: 127977   epsilon: 0.94460356000
12026   steps: 179   lr: 0.0001   evaluation reward: 1.52
episode: 684   score: 0.0   memory length: 128100   epsilon: 0.94436002000
12079   steps: 123   lr: 0.0001   evaluation reward: 1.52
episode: 685   score: 0.0   memory length: 128223   epsilon: 0.94411648000
12132   steps: 123   lr: 0.0001   evaluation reward: 1.52
episode: 686   score: 1.0   memory length: 128373   epsilon: 0.94381948000
12196   steps: 150   lr: 0.0001   evaluation reward: 1.52
episode: 687   score: 0.0   memory length: 128496   epsilon: 0.94357594000
12249   steps: 123   lr: 0.0001   evaluation reward: 1.5
episode: 688   score: 0.0   memory length: 128619   epsilon: 0.94333240000
12302   steps: 123   lr: 0.0001   evaluation reward: 1.5
episode: 689   score: 2.0   memory length: 128836   epsilon: 0.94290274000
12395   steps: 217   lr: 0.0001   evaluation reward: 1.49
episode: 690   score: 0.0   memory length: 128958   epsilon: 0.94266118000
12448   steps: 122   lr: 0.0001   evaluation reward: 1.48
episode: 691   score: 1.0   memory length: 129109   epsilon: 0.94236220000
12513   steps: 151   lr: 0.0001   evaluation reward: 1.47
episode: 692   score: 4.0   memory length: 129408   epsilon: 0.94177018000
12641   steps: 299   lr: 0.0001   evaluation reward: 1.5
episode: 693   score: 1.0   memory length: 129559   epsilon: 0.94147120000
12706   steps: 151   lr: 0.0001   evaluation reward: 1.49
episode: 694   score: 0.0   memory length: 129681   epsilon: 0.94122964000
12758   steps: 122   lr: 0.0001   evaluation reward: 1.49
episode: 695   score: 2.0   memory length: 129878   epsilon: 0.94083958000
12843   steps: 197   lr: 0.0001   evaluation reward: 1.48
episode: 696   score: 0.0   memory length: 130001   epsilon: 0.94059604000
12896   steps: 123   lr: 0.0001   evaluation reward: 1.47
episode: 697   score: 1.0   memory length: 130169   epsilon: 0.94026340000
12968   steps: 168   lr: 0.0001   evaluation reward: 1.47
episode: 698   score: 0.0   memory length: 130292   epsilon: 0.94001986000
13021   steps: 123   lr: 0.0001   evaluation reward: 1.47
episode: 699   score: 2.0   memory length: 130490   epsilon: 0.93962782000
13106   steps: 198   lr: 0.0001   evaluation reward: 1.49
episode: 700   score: 2.0   memory length: 130709   epsilon: 0.93919420000
132   steps: 219   lr: 0.0001   evaluation reward: 1.51
episode: 701   score: 3.0   memory length: 130956   epsilon: 0.93870514000
13307   steps: 247   lr: 0.0001   evaluation reward: 1.5
episode: 702   score: 0.0   memory length: 131078   epsilon: 0.93846358000
```



```
13359  steps: 122  lr: 0.0001  evaluation reward: 1.47
episode: 703  score: 1.0  memory length: 131246  epsilon: 0.93813094000
13431  steps: 168  lr: 0.0001  evaluation reward: 1.46
episode: 704  score: 2.0  memory length: 131443  epsilon: 0.93774088000
13516  steps: 197  lr: 0.0001  evaluation reward: 1.48
episode: 705  score: 1.0  memory length: 131612  epsilon: 0.93740626000
13588  steps: 169  lr: 0.0001  evaluation reward: 1.48
episode: 706  score: 1.0  memory length: 131783  epsilon: 0.93706768000
13662  steps: 171  lr: 0.0001  evaluation reward: 1.47
episode: 707  score: 2.0  memory length: 132002  epsilon: 0.93663406000
13756  steps: 219  lr: 0.0001  evaluation reward: 1.47
episode: 708  score: 2.0  memory length: 132200  epsilon: 0.93624202000
13841  steps: 198  lr: 0.0001  evaluation reward: 1.47
episode: 709  score: 0.0  memory length: 132323  epsilon: 0.93599848000
13894  steps: 123  lr: 0.0001  evaluation reward: 1.44
episode: 710  score: 1.0  memory length: 132492  epsilon: 0.93566386000
13967  steps: 169  lr: 0.0001  evaluation reward: 1.45
episode: 711  score: 0.0  memory length: 132615  epsilon: 0.93542032000
1402  steps: 123  lr: 0.0001  evaluation reward: 1.44
episode: 712  score: 0.0  memory length: 132738  epsilon: 0.93517678000
14072  steps: 123  lr: 0.0001  evaluation reward: 1.41
episode: 713  score: 4.0  memory length: 133004  epsilon: 0.93465010000
14187  steps: 266  lr: 0.0001  evaluation reward: 1.45
episode: 714  score: 0.0  memory length: 133127  epsilon: 0.93440656000
1424  steps: 123  lr: 0.0001  evaluation reward: 1.42
episode: 715  score: 1.0  memory length: 133295  epsilon: 0.93407392000
14312  steps: 168  lr: 0.0001  evaluation reward: 1.43
episode: 716  score: 4.0  memory length: 133572  epsilon: 0.93352546000
14431  steps: 277  lr: 0.0001  evaluation reward: 1.46
episode: 717  score: 0.0  memory length: 133695  epsilon: 0.93328192000
14484  steps: 123  lr: 0.0001  evaluation reward: 1.44
episode: 718  score: 1.0  memory length: 133864  epsilon: 0.93294730000
14556  steps: 169  lr: 0.0001  evaluation reward: 1.43
episode: 719  score: 3.0  memory length: 134114  epsilon: 0.93245230000
14664  steps: 250  lr: 0.0001  evaluation reward: 1.43
episode: 720  score: 0.0  memory length: 134237  epsilon: 0.93220876000
14717  steps: 123  lr: 0.0001  evaluation reward: 1.41
episode: 721  score: 4.0  memory length: 134532  epsilon: 0.93162466000
14844  steps: 295  lr: 0.0001  evaluation reward: 1.45
episode: 722  score: 1.0  memory length: 134701  epsilon: 0.93129004000
14916  steps: 169  lr: 0.0001  evaluation reward: 1.45
episode: 723  score: 2.0  memory length: 134920  epsilon: 0.93085642000
1501  steps: 219  lr: 0.0001  evaluation reward: 1.47
episode: 724  score: 1.0  memory length: 135071  epsilon: 0.93055744000
15075  steps: 151  lr: 0.0001  evaluation reward: 1.48
episode: 725  score: 2.0  memory length: 135269  epsilon: 0.93016540000
1516  steps: 198  lr: 0.0001  evaluation reward: 1.49
episode: 726  score: 2.0  memory length: 135467  epsilon: 0.92977336000
15246  steps: 198  lr: 0.0001  evaluation reward: 1.5
episode: 727  score: 2.0  memory length: 135684  epsilon: 0.92934370000
15339  steps: 217  lr: 0.0001  evaluation reward: 1.5
episode: 728  score: 3.0  memory length: 135931  epsilon: 0.92885464000
15445  steps: 247  lr: 0.0001  evaluation reward: 1.51
```

```
episode: 729  score: 2.0  memory length: 136111  epsilon: 0.92849824000
15522  steps: 180  lr: 0.0001  evaluation reward: 1.53
episode: 730  score: 1.0  memory length: 136280  epsilon: 0.92816362000
15595  steps: 169  lr: 0.0001  evaluation reward: 1.52
episode: 731  score: 1.0  memory length: 136449  epsilon: 0.92782900000
15668  steps: 169  lr: 0.0001  evaluation reward: 1.53
episode: 732  score: 0.0  memory length: 136572  epsilon: 0.92758546000
1572  steps: 123  lr: 0.0001  evaluation reward: 1.53
episode: 733  score: 4.0  memory length: 136867  epsilon: 0.92700136000
15847  steps: 295  lr: 0.0001  evaluation reward: 1.55
episode: 734  score: 0.0  memory length: 136990  epsilon: 0.92675782000
159  steps: 123  lr: 0.0001  evaluation reward: 1.55
episode: 735  score: 1.0  memory length: 137160  epsilon: 0.92642122000
15973  steps: 170  lr: 0.0001  evaluation reward: 1.54
episode: 736  score: 1.0  memory length: 137311  epsilon: 0.92612224000
16038  steps: 151  lr: 0.0001  evaluation reward: 1.55
episode: 737  score: 0.0  memory length: 137434  epsilon: 0.92587870000
16091  steps: 123  lr: 0.0001  evaluation reward: 1.53
episode: 738  score: 1.0  memory length: 137606  epsilon: 0.92553814000
16165  steps: 172  lr: 0.0001  evaluation reward: 1.51
episode: 739  score: 4.0  memory length: 137919  epsilon: 0.92491840000
163  steps: 313  lr: 0.0001  evaluation reward: 1.53
episode: 740  score: 0.0  memory length: 138042  epsilon: 0.92467486000
16352  steps: 123  lr: 0.0001  evaluation reward: 1.52
episode: 741  score: 1.0  memory length: 138193  epsilon: 0.92437588000
16417  steps: 151  lr: 0.0001  evaluation reward: 1.53
episode: 742  score: 1.0  memory length: 138362  epsilon: 0.92404126000
1649  steps: 169  lr: 0.0001  evaluation reward: 1.53
episode: 743  score: 2.0  memory length: 138579  epsilon: 0.92361160000
16583  steps: 217  lr: 0.0001  evaluation reward: 1.53
episode: 744  score: 0.0  memory length: 138701  epsilon: 0.92337004000
16636  steps: 122  lr: 0.0001  evaluation reward: 1.47
episode: 745  score: 8.0  memory length: 139030  epsilon: 0.92271862000
16777  steps: 329  lr: 0.0001  evaluation reward: 1.54
episode: 746  score: 1.0  memory length: 139181  epsilon: 0.92241964000
16842  steps: 151  lr: 0.0001  evaluation reward: 1.52
episode: 747  score: 1.0  memory length: 139333  epsilon: 0.92211868000
16907  steps: 152  lr: 0.0001  evaluation reward: 1.53
episode: 748  score: 0.0  memory length: 139455  epsilon: 0.92187712000
1696  steps: 122  lr: 0.0001  evaluation reward: 1.53
episode: 749  score: 1.0  memory length: 139627  epsilon: 0.92153656000
17034  steps: 172  lr: 0.0001  evaluation reward: 1.51
episode: 750  score: 3.0  memory length: 139852  epsilon: 0.92109106000
1713  steps: 225  lr: 0.0001  evaluation reward: 1.54
episode: 751  score: 1.0  memory length: 140023  epsilon: 0.92075248000
17204  steps: 171  lr: 0.0001  evaluation reward: 1.53
episode: 752  score: 1.0  memory length: 140173  epsilon: 0.92045548000
17268  steps: 150  lr: 0.0001  evaluation reward: 1.52
episode: 753  score: 0.0  memory length: 140296  epsilon: 0.92021194000
17321  steps: 123  lr: 0.0001  evaluation reward: 1.52
episode: 754  score: 1.0  memory length: 140464  epsilon: 0.91987930000
17393  steps: 168  lr: 0.0001  evaluation reward: 1.51
episode: 755  score: 3.0  memory length: 140709  epsilon: 0.91939420000
```

```
17499  steps: 245  lr: 0.0001  evaluation reward: 1.5
episode: 756  score: 1.0  memory length: 140860  epsilon: 0.91909522000
17564  steps: 151  lr: 0.0001  evaluation reward: 1.5
episode: 757  score: 3.0  memory length: 141124  epsilon: 0.91857250000
17677  steps: 264  lr: 0.0001  evaluation reward: 1.49
episode: 758  score: 3.0  memory length: 141370  epsilon: 0.91808542000
17783  steps: 246  lr: 0.0001  evaluation reward: 1.48
episode: 759  score: 1.0  memory length: 141538  epsilon: 0.91775278000
17855  steps: 168  lr: 0.0001  evaluation reward: 1.48
episode: 760  score: 0.0  memory length: 141661  epsilon: 0.91750924000
17908  steps: 123  lr: 0.0001  evaluation reward: 1.43
episode: 761  score: 0.0  memory length: 141784  epsilon: 0.91726570000
17961  steps: 123  lr: 0.0001  evaluation reward: 1.4
episode: 762  score: 3.0  memory length: 142048  epsilon: 0.91674298000
18074  steps: 264  lr: 0.0001  evaluation reward: 1.43
episode: 763  score: 2.0  memory length: 142265  epsilon: 0.91631332000
18168  steps: 217  lr: 0.0001  evaluation reward: 1.43
episode: 764  score: 2.0  memory length: 142483  epsilon: 0.91588168000
18261  steps: 218  lr: 0.0001  evaluation reward: 1.45
episode: 765  score: 3.0  memory length: 142709  epsilon: 0.91543420000
18358  steps: 226  lr: 0.0001  evaluation reward: 1.46
episode: 766  score: 0.0  memory length: 142832  epsilon: 0.91519066000
18411  steps: 123  lr: 0.0001  evaluation reward: 1.44
episode: 767  score: 0.0  memory length: 142954  epsilon: 0.91494910000
18464  steps: 122  lr: 0.0001  evaluation reward: 1.41
episode: 768  score: 0.0  memory length: 143077  epsilon: 0.91470556000
18517  steps: 123  lr: 0.0001  evaluation reward: 1.4
episode: 769  score: 1.0  memory length: 143249  epsilon: 0.91436500000
1859  steps: 172  lr: 0.0001  evaluation reward: 1.4
episode: 770  score: 1.0  memory length: 143418  epsilon: 0.91403038000
18663  steps: 169  lr: 0.0001  evaluation reward: 1.4
episode: 771  score: 2.0  memory length: 143599  epsilon: 0.91367200000
18741  steps: 181  lr: 0.0001  evaluation reward: 1.39
episode: 772  score: 1.0  memory length: 143750  epsilon: 0.91337302000
18806  steps: 151  lr: 0.0001  evaluation reward: 1.39
episode: 773  score: 3.0  memory length: 144016  epsilon: 0.91284634000
1892  steps: 266  lr: 0.0001  evaluation reward: 1.42
episode: 774  score: 2.0  memory length: 144213  epsilon: 0.91245628000
19005  steps: 197  lr: 0.0001  evaluation reward: 1.43
episode: 775  score: 0.0  memory length: 144335  epsilon: 0.91221472000
19057  steps: 122  lr: 0.0001  evaluation reward: 1.43
episode: 776  score: 0.0  memory length: 144458  epsilon: 0.91197118000
1911  steps: 123  lr: 0.0001  evaluation reward: 1.41
episode: 777  score: 0.0  memory length: 144581  epsilon: 0.91172764000
19163  steps: 123  lr: 0.0001  evaluation reward: 1.41
episode: 778  score: 2.0  memory length: 144799  epsilon: 0.91129600000
19257  steps: 218  lr: 0.0001  evaluation reward: 1.42
episode: 779  score: 0.0  memory length: 144922  epsilon: 0.91105246000
1931  steps: 123  lr: 0.0001  evaluation reward: 1.34
episode: 780  score: 0.0  memory length: 145045  epsilon: 0.91080892000
19363  steps: 123  lr: 0.0001  evaluation reward: 1.32
episode: 781  score: 3.0  memory length: 145275  epsilon: 0.91035352000
19461  steps: 230  lr: 0.0001  evaluation reward: 1.34
```

```
episode: 782  score: 2.0  memory length: 145472  epsilon: 0.90996346000
19546  steps: 197  lr: 0.0001  evaluation reward: 1.35
episode: 783  score: 0.0  memory length: 145594  epsilon: 0.90972190000
19598  steps: 122  lr: 0.0001  evaluation reward: 1.33
episode: 784  score: 2.0  memory length: 145810  epsilon: 0.90929422000
19691  steps: 216  lr: 0.0001  evaluation reward: 1.35
episode: 785  score: 0.0  memory length: 145933  epsilon: 0.90905068000
19744  steps: 123  lr: 0.0001  evaluation reward: 1.35
episode: 786  score: 2.0  memory length: 146130  epsilon: 0.90866062000
19829  steps: 197  lr: 0.0001  evaluation reward: 1.36
episode: 787  score: 2.0  memory length: 146328  epsilon: 0.90826858000
19914  steps: 198  lr: 0.0001  evaluation reward: 1.38
episode: 788  score: 2.0  memory length: 146525  epsilon: 0.90787852000
19999  steps: 197  lr: 0.0001  evaluation reward: 1.4
episode: 789  score: 0.0  memory length: 146648  epsilon: 0.90763498000
20052  steps: 123  lr: 0.0001  evaluation reward: 1.38
episode: 790  score: 2.0  memory length: 146866  epsilon: 0.90720334000
20145  steps: 218  lr: 0.0001  evaluation reward: 1.4
episode: 791  score: 2.0  memory length: 147085  epsilon: 0.90676972000
20239  steps: 219  lr: 0.0001  evaluation reward: 1.41
episode: 792  score: 3.0  memory length: 147317  epsilon: 0.90631036000
20339  steps: 232  lr: 0.0001  evaluation reward: 1.4
episode: 793  score: 0.0  memory length: 147439  epsilon: 0.90606880000
20392  steps: 122  lr: 0.0001  evaluation reward: 1.39
episode: 794  score: 4.0  memory length: 147757  epsilon: 0.90543916000
20528  steps: 318  lr: 0.0001  evaluation reward: 1.43
episode: 795  score: 3.0  memory length: 148005  epsilon: 0.90494812000
20635  steps: 248  lr: 0.0001  evaluation reward: 1.44
episode: 796  score: 2.0  memory length: 148203  epsilon: 0.90455608000
2072  steps: 198  lr: 0.0001  evaluation reward: 1.46
episode: 797  score: 0.0  memory length: 148326  epsilon: 0.90431254000
20773  steps: 123  lr: 0.0001  evaluation reward: 1.45
episode: 798  score: 2.0  memory length: 148524  epsilon: 0.90392050000
20858  steps: 198  lr: 0.0001  evaluation reward: 1.47
episode: 799  score: 7.0  memory length: 148770  epsilon: 0.90343342000
20964  steps: 246  lr: 0.0001  evaluation reward: 1.52
episode: 800  score: 2.0  memory length: 148968  epsilon: 0.90304138000
21049  steps: 198  lr: 0.0001  evaluation reward: 1.52
episode: 801  score: 0.0  memory length: 149091  epsilon: 0.90279784000
21102  steps: 123  lr: 0.0001  evaluation reward: 1.49
episode: 802  score: 3.0  memory length: 149338  epsilon: 0.90230878000
21208  steps: 247  lr: 0.0001  evaluation reward: 1.52
episode: 803  score: 0.0  memory length: 149461  epsilon: 0.90206524000
21261  steps: 123  lr: 0.0001  evaluation reward: 1.51
episode: 804  score: 1.0  memory length: 149630  epsilon: 0.90173062000
21333  steps: 169  lr: 0.0001  evaluation reward: 1.5
episode: 805  score: 1.0  memory length: 149781  epsilon: 0.90143164000
21398  steps: 151  lr: 0.0001  evaluation reward: 1.5
episode: 806  score: 3.0  memory length: 150025  epsilon: 0.90094852000
21503  steps: 244  lr: 0.0001  evaluation reward: 1.52
episode: 807  score: 1.0  memory length: 150176  epsilon: 0.90064954000
21568  steps: 151  lr: 0.0001  evaluation reward: 1.51
episode: 808  score: 3.0  memory length: 150443  epsilon: 0.90012088000
```

```
21683  steps: 267  lr: 0.0001  evaluation reward: 1.52
episode: 809  score: 3.0  memory length: 150689  epsilon: 0.89963380000
21789  steps: 246  lr: 0.0001  evaluation reward: 1.55
episode: 810  score: 1.0  memory length: 150840  epsilon: 0.89933482000
21853  steps: 151  lr: 0.0001  evaluation reward: 1.55
episode: 811  score: 1.0  memory length: 150991  epsilon: 0.89903584000
21918  steps: 151  lr: 0.0001  evaluation reward: 1.56
episode: 812  score: 0.0  memory length: 151114  epsilon: 0.89879230000
21971  steps: 123  lr: 0.0001  evaluation reward: 1.56
episode: 813  score: 2.0  memory length: 151311  epsilon: 0.89840224000
22056  steps: 197  lr: 0.0001  evaluation reward: 1.54
episode: 814  score: 2.0  memory length: 151509  epsilon: 0.89801020000
22141  steps: 198  lr: 0.0001  evaluation reward: 1.56
episode: 815  score: 3.0  memory length: 151735  epsilon: 0.89756272000
22238  steps: 226  lr: 0.0001  evaluation reward: 1.58
episode: 816  score: 0.0  memory length: 151857  epsilon: 0.89732116000
22291  steps: 122  lr: 0.0001  evaluation reward: 1.54
episode: 817  score: 4.0  memory length: 152172  epsilon: 0.89669746000
22426  steps: 315  lr: 0.0001  evaluation reward: 1.58
episode: 818  score: 0.0  memory length: 152295  epsilon: 0.89645392000
22479  steps: 123  lr: 0.0001  evaluation reward: 1.57
episode: 819  score: 3.0  memory length: 152543  epsilon: 0.89596288000
22585  steps: 248  lr: 0.0001  evaluation reward: 1.57
episode: 820  score: 2.0  memory length: 152741  epsilon: 0.89557084000
2267  steps: 198  lr: 0.0001  evaluation reward: 1.59
episode: 821  score: 1.0  memory length: 152892  epsilon: 0.89527186000
22735  steps: 151  lr: 0.0001  evaluation reward: 1.56
episode: 822  score: 1.0  memory length: 153043  epsilon: 0.89497288000
228  steps: 151  lr: 0.0001  evaluation reward: 1.56
episode: 823  score: 1.0  memory length: 153194  epsilon: 0.89467390000
22865  steps: 151  lr: 0.0001  evaluation reward: 1.55
episode: 824  score: 3.0  memory length: 153458  epsilon: 0.89415118000
22979  steps: 264  lr: 0.0001  evaluation reward: 1.57
episode: 825  score: 2.0  memory length: 153656  epsilon: 0.89375914000
23064  steps: 198  lr: 0.0001  evaluation reward: 1.57
episode: 826  score: 2.0  memory length: 153837  epsilon: 0.89340076000
23142  steps: 181  lr: 0.0001  evaluation reward: 1.57
episode: 827  score: 2.0  memory length: 154019  epsilon: 0.89304040000
2322  steps: 182  lr: 0.0001  evaluation reward: 1.57
episode: 828  score: 0.0  memory length: 154141  epsilon: 0.89279884000
23272  steps: 122  lr: 0.0001  evaluation reward: 1.54
episode: 829  score: 1.0  memory length: 154311  epsilon: 0.89246224000
23345  steps: 170  lr: 0.0001  evaluation reward: 1.53
episode: 830  score: 1.0  memory length: 154481  epsilon: 0.89212564000
23418  steps: 170  lr: 0.0001  evaluation reward: 1.53
episode: 831  score: 2.0  memory length: 154699  epsilon: 0.89169400000
23512  steps: 218  lr: 0.0001  evaluation reward: 1.54
episode: 832  score: 1.0  memory length: 154871  epsilon: 0.89135344000
23586  steps: 172  lr: 0.0001  evaluation reward: 1.55
episode: 833  score: 2.0  memory length: 155069  epsilon: 0.89096140000
23671  steps: 198  lr: 0.0001  evaluation reward: 1.53
episode: 834  score: 1.0  memory length: 155221  epsilon: 0.89066044000
23737  steps: 152  lr: 0.0001  evaluation reward: 1.54
```

```
episode: 835   score: 1.0   memory length: 155390   epsilon: 0.89032582000
23809   steps: 169   lr: 0.0001   evaluation reward: 1.54
episode: 836   score: 5.0   memory length: 155719   epsilon: 0.88967440000
23951   steps: 329   lr: 0.0001   evaluation reward: 1.58
episode: 837   score: 2.0   memory length: 155917   epsilon: 0.88928236000
24036   steps: 198   lr: 0.0001   evaluation reward: 1.6
episode: 838   score: 6.0   memory length: 156304   epsilon: 0.88851610000
24202   steps: 387   lr: 0.0001   evaluation reward: 1.65
episode: 839   score: 1.0   memory length: 156473   epsilon: 0.88818148000
24275   steps: 169   lr: 0.0001   evaluation reward: 1.62
episode: 840   score: 4.0   memory length: 156730   epsilon: 0.88767262000
24385   steps: 257   lr: 0.0001   evaluation reward: 1.66
episode: 841   score: 2.0   memory length: 156948   epsilon: 0.88724098000
24479   steps: 218   lr: 0.0001   evaluation reward: 1.67
episode: 842   score: 8.0   memory length: 157283   epsilon: 0.88657768000
24623   steps: 335   lr: 0.0001   evaluation reward: 1.74
episode: 843   score: 0.0   memory length: 157406   epsilon: 0.88633414000
24676   steps: 123   lr: 0.0001   evaluation reward: 1.72
episode: 844   score: 2.0   memory length: 157624   epsilon: 0.88590250000
2477   steps: 218   lr: 0.0001   evaluation reward: 1.74
episode: 845   score: 2.0   memory length: 157822   epsilon: 0.88551046000
24855   steps: 198   lr: 0.0001   evaluation reward: 1.68
episode: 846   score: 4.0   memory length: 158098   epsilon: 0.88496398000
24973   steps: 276   lr: 0.0001   evaluation reward: 1.71
episode: 847   score: 2.0   memory length: 158298   epsilon: 0.88456798000
25059   steps: 200   lr: 0.0001   evaluation reward: 1.72
episode: 848   score: 1.0   memory length: 158449   epsilon: 0.88426900000
25124   steps: 151   lr: 0.0001   evaluation reward: 1.73
episode: 849   score: 2.0   memory length: 158647   epsilon: 0.88387696000
25209   steps: 198   lr: 0.0001   evaluation reward: 1.74
episode: 850   score: 1.0   memory length: 158815   epsilon: 0.88354432000
25281   steps: 168   lr: 0.0001   evaluation reward: 1.72
episode: 851   score: 4.0   memory length: 159090   epsilon: 0.88299982000
254   steps: 275   lr: 0.0001   evaluation reward: 1.75
episode: 852   score: 1.0   memory length: 159258   epsilon: 0.88266718000
25472   steps: 168   lr: 0.0001   evaluation reward: 1.75
episode: 853   score: 0.0   memory length: 159381   epsilon: 0.88242364000
25525   steps: 123   lr: 0.0001   evaluation reward: 1.75
episode: 854   score: 3.0   memory length: 159626   epsilon: 0.88193854000
2563   steps: 245   lr: 0.0001   evaluation reward: 1.77
episode: 855   score: 3.0   memory length: 159852   epsilon: 0.88149106000
25727   steps: 226   lr: 0.0001   evaluation reward: 1.77
episode: 856   score: 2.0   memory length: 160050   epsilon: 0.88109902000
25812   steps: 198   lr: 0.0001   evaluation reward: 1.78
episode: 857   score: 3.0   memory length: 160315   epsilon: 0.88057432000
25926   steps: 265   lr: 0.0001   evaluation reward: 1.78
episode: 858   score: 0.0   memory length: 160438   epsilon: 0.88033078000
25979   steps: 123   lr: 0.0001   evaluation reward: 1.75
episode: 859   score: 1.0   memory length: 160607   epsilon: 0.87999616000
26052   steps: 169   lr: 0.0001   evaluation reward: 1.75
episode: 860   score: 1.0   memory length: 160759   epsilon: 0.87969520000
26117   steps: 152   lr: 0.0001   evaluation reward: 1.76
episode: 861   score: 0.0   memory length: 160882   epsilon: 0.87945166000
```

```
2617  steps: 123  lr: 0.0001  evaluation reward: 1.76
episode: 862  score: 0.0  memory length: 161005  epsilon: 0.87920812000
26223  steps: 123  lr: 0.0001  evaluation reward: 1.73
episode: 863  score: 1.0  memory length: 161155  epsilon: 0.87891112000
26287  steps: 150  lr: 0.0001  evaluation reward: 1.72
episode: 864  score: 2.0  memory length: 161373  epsilon: 0.87847948000
26381  steps: 218  lr: 0.0001  evaluation reward: 1.72
episode: 865  score: 1.0  memory length: 161543  epsilon: 0.87814288000
26454  steps: 170  lr: 0.0001  evaluation reward: 1.7
episode: 866  score: 2.0  memory length: 161761  epsilon: 0.87771124000
26548  steps: 218  lr: 0.0001  evaluation reward: 1.72
episode: 867  score: 1.0  memory length: 161932  epsilon: 0.87737266000
26621  steps: 171  lr: 0.0001  evaluation reward: 1.73
episode: 868  score: 2.0  memory length: 162129  epsilon: 0.87698260000
26706  steps: 197  lr: 0.0001  evaluation reward: 1.75
episode: 869  score: 3.0  memory length: 162377  epsilon: 0.87649156000
26812  steps: 248  lr: 0.0001  evaluation reward: 1.77
episode: 870  score: 3.0  memory length: 162642  epsilon: 0.87596686000
26926  steps: 265  lr: 0.0001  evaluation reward: 1.79
episode: 871  score: 0.0  memory length: 162765  epsilon: 0.87572332000
26979  steps: 123  lr: 0.0001  evaluation reward: 1.77
episode: 872  score: 5.0  memory length: 163096  epsilon: 0.87506794000
27122  steps: 331  lr: 0.0001  evaluation reward: 1.81
episode: 873  score: 3.0  memory length: 163342  epsilon: 0.87458086000
27227  steps: 246  lr: 0.0001  evaluation reward: 1.81
episode: 874  score: 2.0  memory length: 163558  epsilon: 0.87415318000
2732  steps: 216  lr: 0.0001  evaluation reward: 1.81
episode: 875  score: 1.0  memory length: 163708  epsilon: 0.87385618000
27385  steps: 150  lr: 0.0001  evaluation reward: 1.82
episode: 876  score: 4.0  memory length: 163974  epsilon: 0.87332950000
27499  steps: 266  lr: 0.0001  evaluation reward: 1.86
episode: 877  score: 2.0  memory length: 164191  epsilon: 0.87289984000
27592  steps: 217  lr: 0.0001  evaluation reward: 1.88
episode: 878  score: 0.0  memory length: 164314  epsilon: 0.87265630000
27645  steps: 123  lr: 0.0001  evaluation reward: 1.86
episode: 879  score: 2.0  memory length: 164494  epsilon: 0.87229990000
27722  steps: 180  lr: 0.0001  evaluation reward: 1.88
episode: 880  score: 2.0  memory length: 164715  epsilon: 0.87186232000
27817  steps: 221  lr: 0.0001  evaluation reward: 1.9
episode: 881  score: 1.0  memory length: 164886  epsilon: 0.87152374000
27891  steps: 171  lr: 0.0001  evaluation reward: 1.88
episode: 882  score: 3.0  memory length: 165098  epsilon: 0.87110398000
27982  steps: 212  lr: 0.0001  evaluation reward: 1.89
episode: 883  score: 0.0  memory length: 165221  epsilon: 0.87086044000
28035  steps: 123  lr: 0.0001  evaluation reward: 1.89
episode: 884  score: 0.0  memory length: 165343  epsilon: 0.87061888000
28087  steps: 122  lr: 0.0001  evaluation reward: 1.87
episode: 885  score: 1.0  memory length: 165515  epsilon: 0.87027832000
28161  steps: 172  lr: 0.0001  evaluation reward: 1.88
episode: 886  score: 0.0  memory length: 165638  epsilon: 0.87003478000
28214  steps: 123  lr: 0.0001  evaluation reward: 1.86
episode: 887  score: 0.0  memory length: 165761  epsilon: 0.86979124000
28267  steps: 123  lr: 0.0001  evaluation reward: 1.84
```

```
episode: 888  score: 3.0  memory length: 166008  epsilon: 0.86930218000
28373  steps: 247  lr: 0.0001  evaluation reward: 1.85
episode: 889  score: 1.0  memory length: 166159  epsilon: 0.86900320000
28438  steps: 151  lr: 0.0001  evaluation reward: 1.86
episode: 890  score: 0.0  memory length: 166282  epsilon: 0.86875966000
28491  steps: 123  lr: 0.0001  evaluation reward: 1.84
episode: 891  score: 2.0  memory length: 166500  epsilon: 0.86832802000
28585  steps: 218  lr: 0.0001  evaluation reward: 1.84
episode: 892  score: 1.0  memory length: 166650  epsilon: 0.86803102000
28649  steps: 150  lr: 0.0001  evaluation reward: 1.82
episode: 893  score: 3.0  memory length: 166875  epsilon: 0.86758552000
28746  steps: 225  lr: 0.0001  evaluation reward: 1.85
episode: 894  score: 1.0  memory length: 167046  epsilon: 0.86724694000
28819  steps: 171  lr: 0.0001  evaluation reward: 1.82
episode: 895  score: 2.0  memory length: 167244  epsilon: 0.86685490000
28904  steps: 198  lr: 0.0001  evaluation reward: 1.81
episode: 896  score: 2.0  memory length: 167460  epsilon: 0.86642722000
28997  steps: 216  lr: 0.0001  evaluation reward: 1.81
episode: 897  score: 1.0  memory length: 167630  epsilon: 0.86609062000
2907  steps: 170  lr: 0.0001  evaluation reward: 1.82
episode: 898  score: 4.0  memory length: 167924  epsilon: 0.86550850000
29197  steps: 294  lr: 0.0001  evaluation reward: 1.84
episode: 899  score: 2.0  memory length: 168142  epsilon: 0.86507686000
2929  steps: 218  lr: 0.0001  evaluation reward: 1.79
episode: 900  score: 3.0  memory length: 168409  epsilon: 0.86454820000
29405  steps: 267  lr: 0.0001  evaluation reward: 1.8
episode: 901  score: 1.0  memory length: 168560  epsilon: 0.86424922000
2947  steps: 151  lr: 0.0001  evaluation reward: 1.81
episode: 902  score: 0.0  memory length: 168683  epsilon: 0.86400568000
29523  steps: 123  lr: 0.0001  evaluation reward: 1.78
episode: 903  score: 2.0  memory length: 168883  epsilon: 0.86360968000
29609  steps: 200  lr: 0.0001  evaluation reward: 1.8
episode: 904  score: 0.0  memory length: 169005  epsilon: 0.86336812000
29661  steps: 122  lr: 0.0001  evaluation reward: 1.79
episode: 905  score: 1.0  memory length: 169156  epsilon: 0.86306914000
29726  steps: 151  lr: 0.0001  evaluation reward: 1.79
episode: 906  score: 0.0  memory length: 169279  epsilon: 0.86282560000
29779  steps: 123  lr: 0.0001  evaluation reward: 1.76
episode: 907  score: 3.0  memory length: 169524  epsilon: 0.86234050000
29885  steps: 245  lr: 0.0001  evaluation reward: 1.78
episode: 908  score: 5.0  memory length: 169849  epsilon: 0.86169700000
30024  steps: 325  lr: 0.0001  evaluation reward: 1.8
episode: 909  score: 1.0  memory length: 170018  epsilon: 0.86136238000
30097  steps: 169  lr: 0.0001  evaluation reward: 1.78
episode: 910  score: 2.0  memory length: 170234  epsilon: 0.86093470000
3019  steps: 216  lr: 0.0001  evaluation reward: 1.79
episode: 911  score: 0.0  memory length: 170357  epsilon: 0.86069116000
30243  steps: 123  lr: 0.0001  evaluation reward: 1.78
episode: 912  score: 2.0  memory length: 170555  epsilon: 0.86029912000
30328  steps: 198  lr: 0.0001  evaluation reward: 1.8
episode: 913  score: 2.0  memory length: 170774  epsilon: 0.85986550000
30422  steps: 219  lr: 0.0001  evaluation reward: 1.8
episode: 914  score: 4.0  memory length: 171049  epsilon: 0.85932100000
```



```
3054    steps: 275    lr: 0.0001    evaluation reward: 1.82
episode: 915    score: 1.0    memory length: 171218    epsilon: 0.85898638000
30613   steps: 169    lr: 0.0001    evaluation reward: 1.8
episode: 916    score: 1.0    memory length: 171369    epsilon: 0.85868740000
30678   steps: 151    lr: 0.0001    evaluation reward: 1.81
episode: 917    score: 2.0    memory length: 171588    epsilon: 0.85825378000
30772   steps: 219    lr: 0.0001    evaluation reward: 1.79
episode: 918    score: 1.0    memory length: 171738    epsilon: 0.85795678000
30836   steps: 150    lr: 0.0001    evaluation reward: 1.8
episode: 919    score: 3.0    memory length: 171985    epsilon: 0.85746772000
30942   steps: 247    lr: 0.0001    evaluation reward: 1.8
episode: 920    score: 2.0    memory length: 172183    epsilon: 0.85707568000
31027   steps: 198    lr: 0.0001    evaluation reward: 1.8
episode: 921    score: 3.0    memory length: 172429    epsilon: 0.85658860000
31133   steps: 246    lr: 0.0001    evaluation reward: 1.82
episode: 922    score: 1.0    memory length: 172580    epsilon: 0.85628962000
31198   steps: 151    lr: 0.0001    evaluation reward: 1.82
episode: 923    score: 0.0    memory length: 172703    epsilon: 0.85604608000
31251   steps: 123    lr: 0.0001    evaluation reward: 1.81
episode: 924    score: 2.0    memory length: 172902    epsilon: 0.85565206000
31337   steps: 199    lr: 0.0001    evaluation reward: 1.8
episode: 925    score: 4.0    memory length: 173197    epsilon: 0.85506796000
31463   steps: 295    lr: 0.0001    evaluation reward: 1.82
episode: 926    score: 3.0    memory length: 173443    epsilon: 0.85458088000
31569   steps: 246    lr: 0.0001    evaluation reward: 1.83
episode: 927    score: 3.0    memory length: 173670    epsilon: 0.85413142000
31667   steps: 227    lr: 0.0001    evaluation reward: 1.84
episode: 928    score: 1.0    memory length: 173821    epsilon: 0.85383244000
31732   steps: 151    lr: 0.0001    evaluation reward: 1.85
episode: 929    score: 3.0    memory length: 174065    epsilon: 0.85334932000
31836   steps: 244    lr: 0.0001    evaluation reward: 1.87
episode: 930    score: 0.0    memory length: 174187    epsilon: 0.85310776000
31889   steps: 122    lr: 0.0001    evaluation reward: 1.86
episode: 931    score: 0.0    memory length: 174310    epsilon: 0.85286422000
31942   steps: 123    lr: 0.0001    evaluation reward: 1.84
episode: 932    score: 1.0    memory length: 174482    epsilon: 0.85252366000
32016   steps: 172    lr: 0.0001    evaluation reward: 1.84
episode: 933    score: 6.0    memory length: 174846    epsilon: 0.85180294000
32172   steps: 364    lr: 0.0001    evaluation reward: 1.88
episode: 934    score: 1.0    memory length: 175014    epsilon: 0.85147030000
32244   steps: 168    lr: 0.0001    evaluation reward: 1.88
episode: 935    score: 3.0    memory length: 175262    epsilon: 0.85097926000
32351   steps: 248    lr: 0.0001    evaluation reward: 1.9
episode: 936    score: 2.0    memory length: 175480    epsilon: 0.85054762000
32445   steps: 218    lr: 0.0001    evaluation reward: 1.87
episode: 937    score: 2.0    memory length: 175698    epsilon: 0.85011598000
32538   steps: 218    lr: 0.0001    evaluation reward: 1.87
episode: 938    score: 2.0    memory length: 175896    epsilon: 0.84972394000
32623   steps: 198    lr: 0.0001    evaluation reward: 1.83
episode: 939    score: 2.0    memory length: 176094    epsilon: 0.84933190000
32709   steps: 198    lr: 0.0001    evaluation reward: 1.84
episode: 940    score: 3.0    memory length: 176343    epsilon: 0.84883888000
32816   steps: 249    lr: 0.0001    evaluation reward: 1.83
```

```
episode: 941  score: 3.0  memory length: 176591  epsilon: 0.84834784000
32922  steps: 248  lr: 0.0001  evaluation reward: 1.84
episode: 942  score: 5.0  memory length: 176933  epsilon: 0.84767068000
33069  steps: 342  lr: 0.0001  evaluation reward: 1.81
episode: 943  score: 2.0  memory length: 177151  epsilon: 0.84723904000
33163  steps: 218  lr: 0.0001  evaluation reward: 1.83
episode: 944  score: 2.0  memory length: 177348  epsilon: 0.84684898000
33248  steps: 197  lr: 0.0001  evaluation reward: 1.83
episode: 945  score: 3.0  memory length: 177595  epsilon: 0.84635992000
33354  steps: 247  lr: 0.0001  evaluation reward: 1.84
episode: 946  score: 1.0  memory length: 177746  epsilon: 0.84606094000
33419  steps: 151  lr: 0.0001  evaluation reward: 1.81
episode: 947  score: 1.0  memory length: 177897  epsilon: 0.84576196000
33484  steps: 151  lr: 0.0001  evaluation reward: 1.8
episode: 948  score: 2.0  memory length: 178117  epsilon: 0.84532636000
33578  steps: 220  lr: 0.0001  evaluation reward: 1.81
episode: 949  score: 0.0  memory length: 178240  epsilon: 0.84508282000
33631  steps: 123  lr: 0.0001  evaluation reward: 1.79
episode: 950  score: 2.0  memory length: 178458  epsilon: 0.84465118000
33725  steps: 218  lr: 0.0001  evaluation reward: 1.8
episode: 951  score: 2.0  memory length: 178656  epsilon: 0.84425914000
3381  steps: 198  lr: 0.0001  evaluation reward: 1.78
episode: 952  score: 0.0  memory length: 178779  epsilon: 0.84401560000
33863  steps: 123  lr: 0.0001  evaluation reward: 1.77
episode: 953  score: 0.0  memory length: 178902  epsilon: 0.84377206000
33916  steps: 123  lr: 0.0001  evaluation reward: 1.77
episode: 954  score: 1.0  memory length: 179073  epsilon: 0.84343348000
33989  steps: 171  lr: 0.0001  evaluation reward: 1.75
episode: 955  score: 2.0  memory length: 179270  epsilon: 0.84304342000
34074  steps: 197  lr: 0.0001  evaluation reward: 1.74
episode: 956  score: 1.0  memory length: 179420  epsilon: 0.84274642000
34138  steps: 150  lr: 0.0001  evaluation reward: 1.73
episode: 957  score: 1.0  memory length: 179570  epsilon: 0.84244942000
34203  steps: 150  lr: 0.0001  evaluation reward: 1.71
episode: 958  score: 2.0  memory length: 179770  epsilon: 0.84205342000
34289  steps: 200  lr: 0.0001  evaluation reward: 1.73
episode: 959  score: 2.0  memory length: 179950  epsilon: 0.84169702000
34366  steps: 180  lr: 0.0001  evaluation reward: 1.74
episode: 960  score: 2.0  memory length: 180148  epsilon: 0.84130498000
34451  steps: 198  lr: 0.0001  evaluation reward: 1.75
episode: 961  score: 2.0  memory length: 180366  epsilon: 0.84087334000
34545  steps: 218  lr: 0.0001  evaluation reward: 1.77
episode: 962  score: 2.0  memory length: 180546  epsilon: 0.84051694000
34622  steps: 180  lr: 0.0001  evaluation reward: 1.79
episode: 963  score: 3.0  memory length: 180774  epsilon: 0.84006550000
3472  steps: 228  lr: 0.0001  evaluation reward: 1.81
episode: 964  score: 0.0  memory length: 180897  epsilon: 0.83982196000
34773  steps: 123  lr: 0.0001  evaluation reward: 1.79
episode: 965  score: 3.0  memory length: 181142  epsilon: 0.83933686000
34878  steps: 245  lr: 0.0001  evaluation reward: 1.81
episode: 966  score: 0.0  memory length: 181265  epsilon: 0.83909332000
34931  steps: 123  lr: 0.0001  evaluation reward: 1.79
episode: 967  score: 3.0  memory length: 181491  epsilon: 0.83864584000
```

```
35028  steps: 226  lr: 0.0001  evaluation reward: 1.81
episode: 968  score: 5.0  memory length: 181799  epsilon: 0.83803600000
35161  steps: 308  lr: 0.0001  evaluation reward: 1.84
episode: 969  score: 2.0  memory length: 182017  epsilon: 0.83760436000
35254  steps: 218  lr: 0.0001  evaluation reward: 1.83
episode: 970  score: 3.0  memory length: 182265  epsilon: 0.83711332000
35361  steps: 248  lr: 0.0001  evaluation reward: 1.83
episode: 971  score: 0.0  memory length: 182388  epsilon: 0.83686978000
35414  steps: 123  lr: 0.0001  evaluation reward: 1.83
episode: 972  score: 4.0  memory length: 182680  epsilon: 0.83629162000
3554  steps: 292  lr: 0.0001  evaluation reward: 1.82
episode: 973  score: 5.0  memory length: 183014  epsilon: 0.83563030000
35683  steps: 334  lr: 0.0001  evaluation reward: 1.84
episode: 974  score: 3.0  memory length: 183260  epsilon: 0.83514322000
35789  steps: 246  lr: 0.0001  evaluation reward: 1.85
episode: 975  score: 2.0  memory length: 183458  epsilon: 0.83475118000
35874  steps: 198  lr: 0.0001  evaluation reward: 1.86
episode: 976  score: 3.0  memory length: 183684  epsilon: 0.83430370000
35971  steps: 226  lr: 0.0001  evaluation reward: 1.85
episode: 977  score: 3.0  memory length: 183928  epsilon: 0.83382058000
36076  steps: 244  lr: 0.0001  evaluation reward: 1.86
episode: 978  score: 3.0  memory length: 184173  epsilon: 0.83333548000
36181  steps: 245  lr: 0.0001  evaluation reward: 1.89
episode: 979  score: 4.0  memory length: 184449  epsilon: 0.83278900000
363  steps: 276  lr: 0.0001  evaluation reward: 1.91
episode: 980  score: 3.0  memory length: 184675  epsilon: 0.83234152000
36397  steps: 226  lr: 0.0001  evaluation reward: 1.92
episode: 981  score: 4.0  memory length: 184971  epsilon: 0.83175544000
36524  steps: 296  lr: 0.0001  evaluation reward: 1.95
episode: 982  score: 6.0  memory length: 185348  epsilon: 0.83100898000
36686  steps: 377  lr: 0.0001  evaluation reward: 1.98
episode: 983  score: 3.0  memory length: 185596  epsilon: 0.83051794000
36793  steps: 248  lr: 0.0001  evaluation reward: 2.01
episode: 984  score: 3.0  memory length: 185826  epsilon: 0.83006254000
36892  steps: 230  lr: 0.0001  evaluation reward: 2.04
episode: 985  score: 3.0  memory length: 186071  epsilon: 0.82957744000
36997  steps: 245  lr: 0.0001  evaluation reward: 2.06
episode: 986  score: 4.0  memory length: 186387  epsilon: 0.82895176000
37133  steps: 316  lr: 0.0001  evaluation reward: 2.1
episode: 987  score: 1.0  memory length: 186538  epsilon: 0.82865278000
37198  steps: 151  lr: 0.0001  evaluation reward: 2.11
episode: 988  score: 2.0  memory length: 186756  epsilon: 0.82822114000
37292  steps: 218  lr: 0.0001  evaluation reward: 2.1
episode: 989  score: 4.0  memory length: 187030  epsilon: 0.82767862000
37409  steps: 274  lr: 0.0001  evaluation reward: 2.13
episode: 990  score: 3.0  memory length: 187277  epsilon: 0.82718956000
37515  steps: 247  lr: 0.0001  evaluation reward: 2.16
episode: 991  score: 2.0  memory length: 187460  epsilon: 0.82682722000
37594  steps: 183  lr: 0.0001  evaluation reward: 2.16
episode: 992  score: 3.0  memory length: 187685  epsilon: 0.82638172000
37691  steps: 225  lr: 0.0001  evaluation reward: 2.18
episode: 993  score: 1.0  memory length: 187854  epsilon: 0.82604710000
37763  steps: 169  lr: 0.0001  evaluation reward: 2.16
```

```
episode: 994   score: 6.0   memory length: 188208   epsilon: 0.82534618000
37916   steps: 354   lr: 0.0001   evaluation reward: 2.21
episode: 995   score: 0.0   memory length: 188331   epsilon: 0.82510264000
37969   steps: 123   lr: 0.0001   evaluation reward: 2.19
episode: 996   score: 0.0   memory length: 188453   epsilon: 0.82486108000
38021   steps: 122   lr: 0.0001   evaluation reward: 2.17
episode: 997   score: 5.0   memory length: 188782   epsilon: 0.82420966000
38162   steps: 329   lr: 0.0001   evaluation reward: 2.21
episode: 998   score: 2.0   memory length: 189002   epsilon: 0.82377406000
38257   steps: 220   lr: 0.0001   evaluation reward: 2.19
episode: 999   score: 3.0   memory length: 189213   epsilon: 0.82335628000
38348   steps: 211   lr: 0.0001   evaluation reward: 2.2
episode: 1000  score: 5.0   memory length: 189555   epsilon: 0.82267912000
038495  steps: 342   lr: 0.0001   evaluation reward: 2.22
episode: 1001  score: 2.0   memory length: 189773   epsilon: 0.82224748000
038588  steps: 218   lr: 0.0001   evaluation reward: 2.23
episode: 1002  score: 3.0   memory length: 190001   epsilon: 0.82179604000
038686  steps: 228   lr: 0.0001   evaluation reward: 2.26
episode: 1003  score: 2.0   memory length: 190181   epsilon: 0.82143964000
038764  steps: 180   lr: 0.0001   evaluation reward: 2.26
episode: 1004  score: 1.0   memory length: 190350   epsilon: 0.82110502000
038836  steps: 169   lr: 0.0001   evaluation reward: 2.27
episode: 1005  score: 2.0   memory length: 190550   epsilon: 0.82070902000
038922  steps: 200   lr: 0.0001   evaluation reward: 2.28
episode: 1006  score: 2.0   memory length: 190768   epsilon: 0.82027738000
039016  steps: 218   lr: 0.0001   evaluation reward: 2.3
episode: 1007  score: 0.0   memory length: 190891   epsilon: 0.82003384000
039069  steps: 123   lr: 0.0001   evaluation reward: 2.27
episode: 1008  score: 2.0   memory length: 191109   epsilon: 0.81960220000
039163  steps: 218   lr: 0.0001   evaluation reward: 2.24
episode: 1009  score: 3.0   memory length: 191338   epsilon: 0.81914878000
039261  steps: 229   lr: 0.0001   evaluation reward: 2.26
episode: 1010  score: 7.0   memory length: 191772   epsilon: 0.81828946000
039448  steps: 434   lr: 0.0001   evaluation reward: 2.31
episode: 1011  score: 1.0   memory length: 191923   epsilon: 0.81799048000
039512  steps: 151   lr: 0.0001   evaluation reward: 2.32
episode: 1012  score: 3.0   memory length: 192172   epsilon: 0.81749746000
03962  steps: 249   lr: 0.0001   evaluation reward: 2.33
episode: 1013  score: 3.0   memory length: 192397   epsilon: 0.81705196000
039716  steps: 225   lr: 0.0001   evaluation reward: 2.34
episode: 1014  score: 4.0   memory length: 192692   epsilon: 0.81646786000
039843  steps: 295   lr: 0.0001   evaluation reward: 2.34
episode: 1015  score: 1.0   memory length: 192843   epsilon: 0.81616888000
039908  steps: 151   lr: 0.0001   evaluation reward: 2.34
episode: 1016  score: 5.0   memory length: 193149   epsilon: 0.81556300000
04004  steps: 306   lr: 0.0001   evaluation reward: 2.38
episode: 1017  score: 2.0   memory length: 193367   epsilon: 0.81513136000
040133  steps: 218   lr: 0.0001   evaluation reward: 2.38
episode: 1018  score: 1.0   memory length: 193518   epsilon: 0.81483238000
040198  steps: 151   lr: 0.0001   evaluation reward: 2.38
episode: 1019  score: 3.0   memory length: 193748   epsilon: 0.81437698000
040297  steps: 230   lr: 0.0001   evaluation reward: 2.38
episode: 1020  score: 0.0   memory length: 193871   epsilon: 0.81413344000
```

```
04035    steps: 123    lr: 0.0001    evaluation reward: 2.36
episode: 1021    score: 2.0    memory length: 194071    epsilon: 0.8137374400
040436    steps: 200    lr: 0.0001    evaluation reward: 2.35
episode: 1022    score: 1.0    memory length: 194222    epsilon: 0.8134384600
040501    steps: 151    lr: 0.0001    evaluation reward: 2.35
episode: 1023    score: 3.0    memory length: 194490    epsilon: 0.8129078200
040616    steps: 268    lr: 0.0001    evaluation reward: 2.38
episode: 1024    score: 0.0    memory length: 194613    epsilon: 0.8126642800
040669    steps: 123    lr: 0.0001    evaluation reward: 2.36
episode: 1025    score: 2.0    memory length: 194832    epsilon: 0.8122306600
040763    steps: 219    lr: 0.0001    evaluation reward: 2.34
episode: 1026    score: 3.0    memory length: 195058    epsilon: 0.8117831800
04086    steps: 226    lr: 0.0001    evaluation reward: 2.34
episode: 1027    score: 3.0    memory length: 195285    epsilon: 0.8113337200
040958    steps: 227    lr: 0.0001    evaluation reward: 2.34
episode: 1028    score: 0.0    memory length: 195408    epsilon: 0.8110901800
04101    steps: 123    lr: 0.0001    evaluation reward: 2.33
episode: 1029    score: 1.0    memory length: 195576    epsilon: 0.8107575400
041083    steps: 168    lr: 0.0001    evaluation reward: 2.31
episode: 1030    score: 1.0    memory length: 195745    epsilon: 0.8104229200
041155    steps: 169    lr: 0.0001    evaluation reward: 2.32
episode: 1031    score: 2.0    memory length: 195943    epsilon: 0.8100308800
04124    steps: 198    lr: 0.0001    evaluation reward: 2.34
episode: 1032    score: 3.0    memory length: 196191    epsilon: 0.8095398400
041347    steps: 248    lr: 0.0001    evaluation reward: 2.36
episode: 1033    score: 4.0    memory length: 196487    epsilon: 0.8089537600
041474    steps: 296    lr: 0.0001    evaluation reward: 2.34
episode: 1034    score: 3.0    memory length: 196712    epsilon: 0.8085082600
041571    steps: 225    lr: 0.0001    evaluation reward: 2.36
episode: 1035    score: 4.0    memory length: 196989    epsilon: 0.8079598000
04169    steps: 277    lr: 0.0001    evaluation reward: 2.37
episode: 1036    score: 1.0    memory length: 197140    epsilon: 0.8076608200
041755    steps: 151    lr: 0.0001    evaluation reward: 2.36
episode: 1037    score: 1.0    memory length: 197309    epsilon: 0.8073262000
041828    steps: 169    lr: 0.0001    evaluation reward: 2.35
episode: 1038    score: 2.0    memory length: 197506    epsilon: 0.8069361400
041912    steps: 197    lr: 0.0001    evaluation reward: 2.35
episode: 1039    score: 4.0    memory length: 197765    epsilon: 0.8064233200
042024    steps: 259    lr: 0.0001    evaluation reward: 2.37
episode: 1040    score: 4.0    memory length: 198019    epsilon: 0.8059204000
042133    steps: 254    lr: 0.0001    evaluation reward: 2.38
episode: 1041    score: 3.0    memory length: 198282    epsilon: 0.8053996600
042246    steps: 263    lr: 0.0001    evaluation reward: 2.38
episode: 1042    score: 1.0    memory length: 198451    epsilon: 0.8050650400
042318    steps: 169    lr: 0.0001    evaluation reward: 2.34
episode: 1043    score: 2.0    memory length: 198649    epsilon: 0.8046730000
042404    steps: 198    lr: 0.0001    evaluation reward: 2.34
episode: 1044    score: 7.0    memory length: 199040    epsilon: 0.8038988200
042572    steps: 391    lr: 0.0001    evaluation reward: 2.39
episode: 1045    score: 2.0    memory length: 199240    epsilon: 0.8035028200
042658    steps: 200    lr: 0.0001    evaluation reward: 2.38
episode: 1046    score: 4.0    memory length: 199521    epsilon: 0.8029464400
042778    steps: 281    lr: 0.0001    evaluation reward: 2.41
```

```
episode: 1047  score: 2.0  memory length: 199720  epsilon: 0.8025524200
042864  steps: 199  lr: 0.0001  evaluation reward: 2.42
episode: 1048  score: 4.0  memory length: 199998  epsilon: 0.8020019800
042983  steps: 278  lr: 0.0001  evaluation reward: 2.44
episode: 1049  score: 3.0  memory length: 200244  epsilon: 0.8015149000
043089  steps: 246  lr: 4e-05  evaluation reward: 2.47
episode: 1050  score: 5.0  memory length: 200569  epsilon: 0.8008714000
043229  steps: 325  lr: 4e-05  evaluation reward: 2.5
episode: 1051  score: 5.0  memory length: 200936  epsilon: 0.8001447400
043387  steps: 367  lr: 4e-05  evaluation reward: 2.53
episode: 1052  score: 2.0  memory length: 201134  epsilon: 0.7997527000
043472  steps: 198  lr: 4e-05  evaluation reward: 2.55
episode: 1053  score: 2.0  memory length: 201353  epsilon: 0.7993190800
043566  steps: 219  lr: 4e-05  evaluation reward: 2.57
episode: 1054  score: 2.0  memory length: 201552  epsilon: 0.7989250600
043651  steps: 199  lr: 4e-05  evaluation reward: 2.58
episode: 1055  score: 7.0  memory length: 201923  epsilon: 0.7981904800
043811  steps: 371  lr: 4e-05  evaluation reward: 2.63
episode: 1056  score: 3.0  memory length: 202148  epsilon: 0.7977449800
043908  steps: 225  lr: 4e-05  evaluation reward: 2.65
episode: 1057  score: 3.0  memory length: 202379  epsilon: 0.7972876000
044007  steps: 231  lr: 4e-05  evaluation reward: 2.67
episode: 1058  score: 5.0  memory length: 202673  epsilon: 0.7967054800
044133  steps: 294  lr: 4e-05  evaluation reward: 2.7
episode: 1059  score: 0.0  memory length: 202795  epsilon: 0.7964639200
044186  steps: 122  lr: 4e-05  evaluation reward: 2.68
episode: 1060  score: 2.0  memory length: 203011  epsilon: 0.7960362400
044279  steps: 216  lr: 4e-05  evaluation reward: 2.68
episode: 1061  score: 2.0  memory length: 203228  epsilon: 0.7956065800
044372  steps: 217  lr: 4e-05  evaluation reward: 2.68
episode: 1062  score: 3.0  memory length: 203453  epsilon: 0.7951610800
044469  steps: 225  lr: 4e-05  evaluation reward: 2.69
episode: 1063  score: 3.0  memory length: 203683  epsilon: 0.7947056800
044567  steps: 230  lr: 4e-05  evaluation reward: 2.69
episode: 1064  score: 1.0  memory length: 203834  epsilon: 0.7944067000
044632  steps: 151  lr: 4e-05  evaluation reward: 2.7
episode: 1065  score: 1.0  memory length: 203985  epsilon: 0.7941077200
044697  steps: 151  lr: 4e-05  evaluation reward: 2.68
episode: 1066  score: 2.0  memory length: 204183  epsilon: 0.7937156800
044782  steps: 198  lr: 4e-05  evaluation reward: 2.7
episode: 1067  score: 4.0  memory length: 204459  epsilon: 0.7931692000
044901  steps: 276  lr: 4e-05  evaluation reward: 2.71
episode: 1068  score: 4.0  memory length: 204775  epsilon: 0.7925435200
045037  steps: 316  lr: 4e-05  evaluation reward: 2.7
episode: 1069  score: 2.0  memory length: 204975  epsilon: 0.7921475200
045123  steps: 200  lr: 4e-05  evaluation reward: 2.7
episode: 1070  score: 5.0  memory length: 205315  epsilon: 0.7914743200
045269  steps: 340  lr: 4e-05  evaluation reward: 2.72
episode: 1071  score: 1.0  memory length: 205484  epsilon: 0.7911397000
045342  steps: 169  lr: 4e-05  evaluation reward: 2.73
episode: 1072  score: 2.0  memory length: 205664  epsilon: 0.7907833000
045419  steps: 180  lr: 4e-05  evaluation reward: 2.71
episode: 1073  score: 3.0  memory length: 205927  epsilon: 0.7902625600
```

```
045532    steps: 263    lr: 4e-05    evaluation reward: 2.69
episode: 1074    score: 2.0    memory length: 206145    epsilon: 0.7898309200
045626    steps: 218    lr: 4e-05    evaluation reward: 2.68
episode: 1075    score: 2.0    memory length: 206325    epsilon: 0.7894745200
045703    steps: 180    lr: 4e-05    evaluation reward: 2.68
episode: 1076    score: 1.0    memory length: 206495    epsilon: 0.7891379200
045776    steps: 170    lr: 4e-05    evaluation reward: 2.66
episode: 1077    score: 4.0    memory length: 206770    epsilon: 0.7885934200
045894    steps: 275    lr: 4e-05    evaluation reward: 2.67
episode: 1078    score: 2.0    memory length: 206968    epsilon: 0.7882013800
045979    steps: 198    lr: 4e-05    evaluation reward: 2.66
episode: 1079    score: 1.0    memory length: 207139    epsilon: 0.7878628000
046053    steps: 171    lr: 4e-05    evaluation reward: 2.63
episode: 1080    score: 2.0    memory length: 207319    epsilon: 0.7875064000
04613    steps: 180    lr: 4e-05    evaluation reward: 2.62
episode: 1081    score: 2.0    memory length: 207537    epsilon: 0.7870747600
046224    steps: 218    lr: 4e-05    evaluation reward: 2.6
episode: 1082    score: 5.0    memory length: 207857    epsilon: 0.7864411600
046362    steps: 320    lr: 4e-05    evaluation reward: 2.59
episode: 1083    score: 3.0    memory length: 208085    epsilon: 0.7859897200
04646    steps: 228    lr: 4e-05    evaluation reward: 2.59
episode: 1084    score: 3.0    memory length: 208295    epsilon: 0.7855739200
04655    steps: 210    lr: 4e-05    evaluation reward: 2.59
episode: 1085    score: 2.0    memory length: 208494    epsilon: 0.7851799000
046635    steps: 199    lr: 4e-05    evaluation reward: 2.58
episode: 1086    score: 1.0    memory length: 208645    epsilon: 0.7848809200
0467    steps: 151    lr: 4e-05    evaluation reward: 2.55
episode: 1087    score: 1.0    memory length: 208795    epsilon: 0.7845839200
046765    steps: 150    lr: 4e-05    evaluation reward: 2.55
episode: 1088    score: 5.0    memory length: 209069    epsilon: 0.7840414000
046882    steps: 274    lr: 4e-05    evaluation reward: 2.58
episode: 1089    score: 6.0    memory length: 209410    epsilon: 0.7833662200
047029    steps: 341    lr: 4e-05    evaluation reward: 2.6
episode: 1090    score: 7.0    memory length: 209851    epsilon: 0.7824930400
047219    steps: 441    lr: 4e-05    evaluation reward: 2.64
episode: 1091    score: 5.0    memory length: 210175    epsilon: 0.7818515200
047358    steps: 324    lr: 4e-05    evaluation reward: 2.67
episode: 1092    score: 0.0    memory length: 210298    epsilon: 0.7816079800
047411    steps: 123    lr: 4e-05    evaluation reward: 2.64
episode: 1093    score: 7.0    memory length: 210698    epsilon: 0.7808159800
047583    steps: 400    lr: 4e-05    evaluation reward: 2.7
episode: 1094    score: 1.0    memory length: 210868    epsilon: 0.7804793800
047656    steps: 170    lr: 4e-05    evaluation reward: 2.65
episode: 1095    score: 6.0    memory length: 211218    epsilon: 0.7797863800
047806    steps: 350    lr: 4e-05    evaluation reward: 2.71
episode: 1096    score: 4.0    memory length: 211535    epsilon: 0.7791587200
047942    steps: 317    lr: 4e-05    evaluation reward: 2.75
episode: 1097    score: 3.0    memory length: 211781    epsilon: 0.7786716400
048048    steps: 246    lr: 4e-05    evaluation reward: 2.73
episode: 1098    score: 3.0    memory length: 212009    epsilon: 0.7782202000
048146    steps: 228    lr: 4e-05    evaluation reward: 2.74
episode: 1099    score: 2.0    memory length: 212190    epsilon: 0.7778618200
048224    steps: 181    lr: 4e-05    evaluation reward: 2.73
```

```
episode: 1100  score: 2.0  memory length: 212392  epsilon: 0.7774618600
048311  steps: 202  lr: 4e-05  evaluation reward: 2.7
episode: 1101  score: 4.0  memory length: 212650  epsilon: 0.7769510200
048422  steps: 258  lr: 4e-05  evaluation reward: 2.72
episode: 1102  score: 4.0  memory length: 212908  epsilon: 0.7764401800
048533  steps: 258  lr: 4e-05  evaluation reward: 2.73
episode: 1103  score: 2.0  memory length: 213105  epsilon: 0.7760501200
048617  steps: 197  lr: 4e-05  evaluation reward: 2.73
episode: 1104  score: 3.0  memory length: 213370  epsilon: 0.7755254200
048731  steps: 265  lr: 4e-05  evaluation reward: 2.75
episode: 1105  score: 1.0  memory length: 213520  epsilon: 0.7752284200
048796  steps: 150  lr: 4e-05  evaluation reward: 2.74
episode: 1106  score: 1.0  memory length: 213689  epsilon: 0.7748938000
048868  steps: 169  lr: 4e-05  evaluation reward: 2.73
episode: 1107  score: 2.0  memory length: 213868  epsilon: 0.7745393800
048945  steps: 179  lr: 4e-05  evaluation reward: 2.75
episode: 1108  score: 5.0  memory length: 214217  epsilon: 0.7738483600
049095  steps: 349  lr: 4e-05  evaluation reward: 2.78
episode: 1109  score: 5.0  memory length: 214534  epsilon: 0.7732207000
049232  steps: 317  lr: 4e-05  evaluation reward: 2.8
episode: 1110  score: 4.0  memory length: 214792  epsilon: 0.7727098600
049342  steps: 258  lr: 4e-05  evaluation reward: 2.77
episode: 1111  score: 5.0  memory length: 215117  epsilon: 0.7720663600
049482  steps: 325  lr: 4e-05  evaluation reward: 2.81
episode: 1112  score: 0.0  memory length: 215239  epsilon: 0.7718248000
049535  steps: 122  lr: 4e-05  evaluation reward: 2.78
episode: 1113  score: 7.0  memory length: 215655  epsilon: 0.7710011200
049713  steps: 416  lr: 4e-05  evaluation reward: 2.82
episode: 1114  score: 3.0  memory length: 215902  epsilon: 0.7705120600
04982  steps: 247  lr: 4e-05  evaluation reward: 2.81
episode: 1115  score: 2.0  memory length: 216082  epsilon: 0.7701556600
049897  steps: 180  lr: 4e-05  evaluation reward: 2.82
episode: 1116  score: 2.0  memory length: 216280  epsilon: 0.7697636200
049982  steps: 198  lr: 4e-05  evaluation reward: 2.79
episode: 1117  score: 9.0  memory length: 216773  epsilon: 0.7687874800
050194  steps: 493  lr: 4e-05  evaluation reward: 2.86
episode: 1118  score: 11.0  memory length: 217223  epsilon: 0.767896480
0050387  steps: 450  lr: 4e-05  evaluation reward: 2.96
episode: 1119  score: 3.0  memory length: 217469  epsilon: 0.7674094000
050493  steps: 246  lr: 4e-05  evaluation reward: 2.96
episode: 1120  score: 1.0  memory length: 217619  epsilon: 0.7671124000
050558  steps: 150  lr: 4e-05  evaluation reward: 2.97
episode: 1121  score: 3.0  memory length: 217845  epsilon: 0.7666649200
050655  steps: 226  lr: 4e-05  evaluation reward: 2.98
episode: 1122  score: 4.0  memory length: 218122  epsilon: 0.7661164600
050774  steps: 277  lr: 4e-05  evaluation reward: 3.01
episode: 1123  score: 3.0  memory length: 218349  epsilon: 0.7656670000
050871  steps: 227  lr: 4e-05  evaluation reward: 3.01
episode: 1124  score: 3.0  memory length: 218612  epsilon: 0.7651462600
050984  steps: 263  lr: 4e-05  evaluation reward: 3.04
episode: 1125  score: 2.0  memory length: 218810  epsilon: 0.7647542200
05107  steps: 198  lr: 4e-05  evaluation reward: 3.04
episode: 1126  score: 2.0  memory length: 219028  epsilon: 0.7643225800
```



```
051163    steps: 218    lr: 4e-05    evaluation reward: 3.03
episode: 1127    score: 3.0    memory length: 219280    epsilon: 0.7638236200
051272    steps: 252    lr: 4e-05    evaluation reward: 3.03
episode: 1128    score: 2.0    memory length: 219461    epsilon: 0.7634652400
051349    steps: 181    lr: 4e-05    evaluation reward: 3.05
episode: 1129    score: 5.0    memory length: 219767    epsilon: 0.7628593600
051481    steps: 306    lr: 4e-05    evaluation reward: 3.09
episode: 1130    score: 5.0    memory length: 220113    epsilon: 0.7621742800
05163     steps: 346    lr: 4e-05    evaluation reward: 3.13
episode: 1131    score: 1.0    memory length: 220284    epsilon: 0.7618357000
051703    steps: 171    lr: 4e-05    evaluation reward: 3.12
episode: 1132    score: 2.0    memory length: 220466    epsilon: 0.7614753400
051781    steps: 182    lr: 4e-05    evaluation reward: 3.11
episode: 1133    score: 4.0    memory length: 220763    epsilon: 0.7608872800
051909    steps: 297    lr: 4e-05    evaluation reward: 3.11
episode: 1134    score: 1.0    memory length: 220932    epsilon: 0.7605526600
051982    steps: 169    lr: 4e-05    evaluation reward: 3.09
episode: 1135    score: 3.0    memory length: 221200    epsilon: 0.7600220200
052097    steps: 268    lr: 4e-05    evaluation reward: 3.08
episode: 1136    score: 5.0    memory length: 221546    epsilon: 0.7593369400
052246    steps: 346    lr: 4e-05    evaluation reward: 3.12
episode: 1137    score: 0.0    memory length: 221669    epsilon: 0.7590934000
052298    steps: 123    lr: 4e-05    evaluation reward: 3.11
episode: 1138    score: 5.0    memory length: 222014    epsilon: 0.7584103000
052447    steps: 345    lr: 4e-05    evaluation reward: 3.14
episode: 1139    score: 3.0    memory length: 222259    epsilon: 0.7579252000
052552    steps: 245    lr: 4e-05    evaluation reward: 3.13
episode: 1140    score: 0.0    memory length: 222382    epsilon: 0.7576816600
052605    steps: 123    lr: 4e-05    evaluation reward: 3.09
episode: 1141    score: 6.0    memory length: 222750    epsilon: 0.7569530200
052763    steps: 368    lr: 4e-05    evaluation reward: 3.12
episode: 1142    score: 4.0    memory length: 223026    epsilon: 0.7564065400
052882    steps: 276    lr: 4e-05    evaluation reward: 3.15
episode: 1143    score: 3.0    memory length: 223270    epsilon: 0.7559234200
052987    steps: 244    lr: 4e-05    evaluation reward: 3.16
episode: 1144    score: 2.0    memory length: 223471    epsilon: 0.7555254400
053073    steps: 201    lr: 4e-05    evaluation reward: 3.11
episode: 1145    score: 3.0    memory length: 223718    epsilon: 0.7550363800
053179    steps: 247    lr: 4e-05    evaluation reward: 3.12
episode: 1146    score: 4.0    memory length: 224020    epsilon: 0.7544384200
053309    steps: 302    lr: 4e-05    evaluation reward: 3.12
episode: 1147    score: 3.0    memory length: 224288    epsilon: 0.7539077800
053424    steps: 268    lr: 4e-05    evaluation reward: 3.13
episode: 1148    score: 1.0    memory length: 224440    epsilon: 0.7536068200
05349     steps: 152    lr: 4e-05    evaluation reward: 3.1
episode: 1149    score: 5.0    memory length: 224733    epsilon: 0.7530266800
053615    steps: 293    lr: 4e-05    evaluation reward: 3.12
episode: 1150    score: 4.0    memory length: 225021    epsilon: 0.7524564400
053739    steps: 288    lr: 4e-05    evaluation reward: 3.11
episode: 1151    score: 3.0    memory length: 225288    epsilon: 0.7519277800
053854    steps: 267    lr: 4e-05    evaluation reward: 3.09
episode: 1152    score: 3.0    memory length: 225516    epsilon: 0.7514763400
053952    steps: 228    lr: 4e-05    evaluation reward: 3.1
```

```
episode: 1153  score: 4.0  memory length: 225816  epsilon: 0.7508823400
054081  steps: 300  lr: 4e-05  evaluation reward: 3.12
episode: 1154  score: 2.0  memory length: 226034  epsilon: 0.7504507000
054175  steps: 218  lr: 4e-05  evaluation reward: 3.12
episode: 1155  score: 2.0  memory length: 226234  epsilon: 0.7500547000
054261  steps: 200  lr: 4e-05  evaluation reward: 3.07
episode: 1156  score: 3.0  memory length: 226480  epsilon: 0.7495676200
054366  steps: 246  lr: 4e-05  evaluation reward: 3.07
episode: 1157  score: 4.0  memory length: 226757  epsilon: 0.7490191600
054485  steps: 277  lr: 4e-05  evaluation reward: 3.08
episode: 1158  score: 2.0  memory length: 226955  epsilon: 0.7486271200
054571  steps: 198  lr: 4e-05  evaluation reward: 3.05
episode: 1159  score: 5.0  memory length: 227277  epsilon: 0.7479895600
054709  steps: 322  lr: 4e-05  evaluation reward: 3.1
episode: 1160  score: 3.0  memory length: 227522  epsilon: 0.7475044600
054814  steps: 245  lr: 4e-05  evaluation reward: 3.11
episode: 1161  score: 5.0  memory length: 227848  epsilon: 0.7468589800
054954  steps: 326  lr: 4e-05  evaluation reward: 3.14
episode: 1162  score: 3.0  memory length: 228095  epsilon: 0.7463699200
055061  steps: 247  lr: 4e-05  evaluation reward: 3.14
episode: 1163  score: 1.0  memory length: 228248  epsilon: 0.7460669800
055126  steps: 153  lr: 4e-05  evaluation reward: 3.12
episode: 1164  score: 6.0  memory length: 228563  epsilon: 0.7454432800
055262  steps: 315  lr: 4e-05  evaluation reward: 3.17
episode: 1165  score: 4.0  memory length: 228820  epsilon: 0.7449344200
055372  steps: 257  lr: 4e-05  evaluation reward: 3.2
episode: 1166  score: 5.0  memory length: 229144  epsilon: 0.7442929000
055512  steps: 324  lr: 4e-05  evaluation reward: 3.23
episode: 1167  score: 3.0  memory length: 229373  epsilon: 0.7438394800
05561  steps: 229  lr: 4e-05  evaluation reward: 3.22
episode: 1168  score: 2.0  memory length: 229591  epsilon: 0.7434078400
055704  steps: 218  lr: 4e-05  evaluation reward: 3.2
episode: 1169  score: 3.0  memory length: 229837  epsilon: 0.7429207600
055809  steps: 246  lr: 4e-05  evaluation reward: 3.21
episode: 1170  score: 5.0  memory length: 230143  epsilon: 0.7423148800
055941  steps: 306  lr: 4e-05  evaluation reward: 3.21
episode: 1171  score: 1.0  memory length: 230315  epsilon: 0.7419743200
056015  steps: 172  lr: 4e-05  evaluation reward: 3.21
episode: 1172  score: 2.0  memory length: 230497  epsilon: 0.7416139600
056093  steps: 182  lr: 4e-05  evaluation reward: 3.21
episode: 1173  score: 6.0  memory length: 230827  epsilon: 0.7409605600
056235  steps: 330  lr: 4e-05  evaluation reward: 3.24
episode: 1174  score: 3.0  memory length: 231071  epsilon: 0.7404774400
05634  steps: 244  lr: 4e-05  evaluation reward: 3.25
episode: 1175  score: 4.0  memory length: 231367  epsilon: 0.7398913600
056467  steps: 296  lr: 4e-05  evaluation reward: 3.27
episode: 1176  score: 4.0  memory length: 231661  epsilon: 0.7393092400
056593  steps: 294  lr: 4e-05  evaluation reward: 3.3
episode: 1177  score: 3.0  memory length: 231889  epsilon: 0.7388578000
056691  steps: 228  lr: 4e-05  evaluation reward: 3.29
episode: 1178  score: 4.0  memory length: 232184  epsilon: 0.7382737000
056818  steps: 295  lr: 4e-05  evaluation reward: 3.31
episode: 1179  score: 0.0  memory length: 232306  epsilon: 0.7380321400
```

```
056871 steps: 122 lr: 4e-05 evaluation reward: 3.3
episode: 1180 score: 3.0 memory length: 232572 epsilon: 0.7375054600
056985 steps: 266 lr: 4e-05 evaluation reward: 3.31
episode: 1181 score: 1.0 memory length: 232723 epsilon: 0.7372064800
05705 steps: 151 lr: 4e-05 evaluation reward: 3.3
episode: 1182 score: 2.0 memory length: 232905 epsilon: 0.7368461200
057128 steps: 182 lr: 4e-05 evaluation reward: 3.27
episode: 1183 score: 3.0 memory length: 233151 epsilon: 0.7363590400
057234 steps: 246 lr: 4e-05 evaluation reward: 3.27
episode: 1184 score: 3.0 memory length: 233394 epsilon: 0.7358779000
057338 steps: 243 lr: 4e-05 evaluation reward: 3.27
episode: 1185 score: 0.0 memory length: 233517 epsilon: 0.7356343600
057391 steps: 123 lr: 4e-05 evaluation reward: 3.25
episode: 1186 score: 0.0 memory length: 233640 epsilon: 0.7353908200
057444 steps: 123 lr: 4e-05 evaluation reward: 3.24
episode: 1187 score: 2.0 memory length: 233860 epsilon: 0.7349552200
057539 steps: 220 lr: 4e-05 evaluation reward: 3.25
episode: 1188 score: 1.0 memory length: 234029 epsilon: 0.7346206000
057611 steps: 169 lr: 4e-05 evaluation reward: 3.21
episode: 1189 score: 2.0 memory length: 234227 epsilon: 0.7342285600
057696 steps: 198 lr: 4e-05 evaluation reward: 3.17
episode: 1190 score: 2.0 memory length: 234425 epsilon: 0.7338365200
057781 steps: 198 lr: 4e-05 evaluation reward: 3.12
episode: 1191 score: 2.0 memory length: 234607 epsilon: 0.7334761600
05786 steps: 182 lr: 4e-05 evaluation reward: 3.09
episode: 1192 score: 14.0 memory length: 235087 epsilon: 0.732525760
0058066 steps: 480 lr: 4e-05 evaluation reward: 3.23
episode: 1193 score: 1.0 memory length: 235259 epsilon: 0.7321852000
05814 steps: 172 lr: 4e-05 evaluation reward: 3.17
episode: 1194 score: 3.0 memory length: 235505 epsilon: 0.7316981200
058246 steps: 246 lr: 4e-05 evaluation reward: 3.19
episode: 1195 score: 3.0 memory length: 235752 epsilon: 0.7312090600
058352 steps: 247 lr: 4e-05 evaluation reward: 3.16
episode: 1196 score: 6.0 memory length: 236100 epsilon: 0.7305200200
058501 steps: 348 lr: 4e-05 evaluation reward: 3.18
episode: 1197 score: 4.0 memory length: 236378 epsilon: 0.7299695800
058621 steps: 278 lr: 4e-05 evaluation reward: 3.19
episode: 1198 score: 2.0 memory length: 236578 epsilon: 0.7295735800
058707 steps: 200 lr: 4e-05 evaluation reward: 3.18
episode: 1199 score: 5.0 memory length: 236906 epsilon: 0.7289241400
058848 steps: 328 lr: 4e-05 evaluation reward: 3.21
episode: 1200 score: 0.0 memory length: 237029 epsilon: 0.7286806000
058901 steps: 123 lr: 4e-05 evaluation reward: 3.19
episode: 1201 score: 3.0 memory length: 237275 epsilon: 0.7281935200
059007 steps: 246 lr: 4e-05 evaluation reward: 3.18
episode: 1202 score: 2.0 memory length: 237473 epsilon: 0.7278014800
059092 steps: 198 lr: 4e-05 evaluation reward: 3.16
episode: 1203 score: 4.0 memory length: 237768 epsilon: 0.7272173800
059218 steps: 295 lr: 4e-05 evaluation reward: 3.18
episode: 1204 score: 3.0 memory length: 237999 epsilon: 0.7267600000
059318 steps: 231 lr: 4e-05 evaluation reward: 3.18
episode: 1205 score: 3.0 memory length: 238225 epsilon: 0.7263125200
059415 steps: 226 lr: 4e-05 evaluation reward: 3.2
```

```
episode: 1206  score: 1.0  memory length: 238396  epsilon: 0.7259739400
059488  steps: 171  lr: 4e-05  evaluation reward: 3.2
episode: 1207  score: 3.0  memory length: 238622  epsilon: 0.7255264600
059586  steps: 226  lr: 4e-05  evaluation reward: 3.21
episode: 1208  score: 3.0  memory length: 238869  epsilon: 0.7250374000
059692  steps: 247  lr: 4e-05  evaluation reward: 3.19
episode: 1209  score: 3.0  memory length: 239094  epsilon: 0.7245919000
059788  steps: 225  lr: 4e-05  evaluation reward: 3.17
episode: 1210  score: 5.0  memory length: 239441  epsilon: 0.7239048400
059938  steps: 347  lr: 4e-05  evaluation reward: 3.18
episode: 1211  score: 3.0  memory length: 239707  epsilon: 0.7233781600
060052  steps: 266  lr: 4e-05  evaluation reward: 3.16
episode: 1212  score: 1.0  memory length: 239858  epsilon: 0.7230791800
060117  steps: 151  lr: 4e-05  evaluation reward: 3.17
episode: 1213  score: 3.0  memory length: 240104  epsilon: 0.7225921000
060223  steps: 246  lr: 4e-05  evaluation reward: 3.13
episode: 1214  score: 4.0  memory length: 240383  epsilon: 0.7220396800
060342  steps: 279  lr: 4e-05  evaluation reward: 3.14
episode: 1215  score: 7.0  memory length: 240780  epsilon: 0.7212536200
060513  steps: 397  lr: 4e-05  evaluation reward: 3.19
episode: 1216  score: 4.0  memory length: 241039  epsilon: 0.7207408000
060624  steps: 259  lr: 4e-05  evaluation reward: 3.21
episode: 1217  score: 4.0  memory length: 241333  epsilon: 0.7201586800
060751  steps: 294  lr: 4e-05  evaluation reward: 3.16
episode: 1218  score: 7.0  memory length: 241775  epsilon: 0.7192835200
060941  steps: 442  lr: 4e-05  evaluation reward: 3.12
episode: 1219  score: 3.0  memory length: 242021  epsilon: 0.7187964400
061047  steps: 246  lr: 4e-05  evaluation reward: 3.12
episode: 1220  score: 5.0  memory length: 242345  epsilon: 0.7181549200
061186  steps: 324  lr: 4e-05  evaluation reward: 3.16
episode: 1221  score: 5.0  memory length: 242667  epsilon: 0.7175173600
061324  steps: 322  lr: 4e-05  evaluation reward: 3.18
episode: 1222  score: 0.0  memory length: 242789  epsilon: 0.7172758000
061377  steps: 122  lr: 4e-05  evaluation reward: 3.14
episode: 1223  score: 6.0  memory length: 243161  epsilon: 0.7165392400
061537  steps: 372  lr: 4e-05  evaluation reward: 3.17
episode: 1224  score: 3.0  memory length: 243390  epsilon: 0.7160858200
061635  steps: 229  lr: 4e-05  evaluation reward: 3.17
episode: 1225  score: 0.0  memory length: 243513  epsilon: 0.7158422800
061688  steps: 123  lr: 4e-05  evaluation reward: 3.15
episode: 1226  score: 2.0  memory length: 243711  epsilon: 0.7154502400
061773  steps: 198  lr: 4e-05  evaluation reward: 3.15
episode: 1227  score: 4.0  memory length: 243988  epsilon: 0.7149017800
061892  steps: 277  lr: 4e-05  evaluation reward: 3.16
episode: 1228  score: 4.0  memory length: 244260  epsilon: 0.7143632200
062009  steps: 272  lr: 4e-05  evaluation reward: 3.18
episode: 1229  score: 2.0  memory length: 244476  epsilon: 0.7139355400
062102  steps: 216  lr: 4e-05  evaluation reward: 3.15
episode: 1230  score: 1.0  memory length: 244645  epsilon: 0.7136009200
062174  steps: 169  lr: 4e-05  evaluation reward: 3.11
episode: 1231  score: 3.0  memory length: 244874  epsilon: 0.7131475000
062273  steps: 229  lr: 4e-05  evaluation reward: 3.13
episode: 1232  score: 8.0  memory length: 245273  epsilon: 0.7123574800
```

```
062444 steps: 399 lr: 4e-05 evaluation reward: 3.19
episode: 1233 score: 4.0 memory length: 245535 epsilon: 0.7118387200
062557 steps: 262 lr: 4e-05 evaluation reward: 3.19
episode: 1234 score: 1.0 memory length: 245706 epsilon: 0.7115001400
06263 steps: 171 lr: 4e-05 evaluation reward: 3.19
episode: 1235 score: 3.0 memory length: 245916 epsilon: 0.7110843400
062721 steps: 210 lr: 4e-05 evaluation reward: 3.19
episode: 1236 score: 3.0 memory length: 246164 epsilon: 0.7105933000
062827 steps: 248 lr: 4e-05 evaluation reward: 3.17
episode: 1237 score: 5.0 memory length: 246467 epsilon: 0.7099933600
062958 steps: 303 lr: 4e-05 evaluation reward: 3.22
episode: 1238 score: 6.0 memory length: 246800 epsilon: 0.7093340200
063101 steps: 333 lr: 4e-05 evaluation reward: 3.23
episode: 1239 score: 0.0 memory length: 246922 epsilon: 0.7090924600
063153 steps: 122 lr: 4e-05 evaluation reward: 3.2
episode: 1240 score: 1.0 memory length: 247091 epsilon: 0.7087578400
063226 steps: 169 lr: 4e-05 evaluation reward: 3.21
episode: 1241 score: 2.0 memory length: 247289 epsilon: 0.7083658000
063311 steps: 198 lr: 4e-05 evaluation reward: 3.17
episode: 1242 score: 2.0 memory length: 247486 epsilon: 0.7079757400
063396 steps: 197 lr: 4e-05 evaluation reward: 3.15
episode: 1243 score: 3.0 memory length: 247732 epsilon: 0.7074886600
063501 steps: 246 lr: 4e-05 evaluation reward: 3.15
episode: 1244 score: 3.0 memory length: 247961 epsilon: 0.7070352400
0636 steps: 229 lr: 4e-05 evaluation reward: 3.16
episode: 1245 score: 0.0 memory length: 248084 epsilon: 0.7067917000
063653 steps: 123 lr: 4e-05 evaluation reward: 3.13
episode: 1246 score: 2.0 memory length: 248281 epsilon: 0.7064016400
063737 steps: 197 lr: 4e-05 evaluation reward: 3.11
episode: 1247 score: 1.0 memory length: 248452 epsilon: 0.7060630600
063811 steps: 171 lr: 4e-05 evaluation reward: 3.09
episode: 1248 score: 4.0 memory length: 248732 epsilon: 0.7055086600
063931 steps: 280 lr: 4e-05 evaluation reward: 3.12
episode: 1249 score: 2.0 memory length: 248929 epsilon: 0.7051186000
064016 steps: 197 lr: 4e-05 evaluation reward: 3.09
episode: 1250 score: 4.0 memory length: 249208 epsilon: 0.7045661800
064136 steps: 279 lr: 4e-05 evaluation reward: 3.09
episode: 1251 score: 5.0 memory length: 249552 epsilon: 0.7038850600
064284 steps: 344 lr: 4e-05 evaluation reward: 3.11
episode: 1252 score: 4.0 memory length: 249803 epsilon: 0.7033880800
064392 steps: 251 lr: 4e-05 evaluation reward: 3.12
episode: 1253 score: 4.0 memory length: 250077 epsilon: 0.7028455600
064509 steps: 274 lr: 4e-05 evaluation reward: 3.12
episode: 1254 score: 0.0 memory length: 250200 epsilon: 0.7026020200
064562 steps: 123 lr: 4e-05 evaluation reward: 3.1
episode: 1255 score: 0.0 memory length: 250323 epsilon: 0.7023584800
064615 steps: 123 lr: 4e-05 evaluation reward: 3.08
episode: 1256 score: 3.0 memory length: 250532 epsilon: 0.7019446600
064705 steps: 209 lr: 4e-05 evaluation reward: 3.08
episode: 1257 score: 3.0 memory length: 250761 epsilon: 0.7014912400
064803 steps: 229 lr: 4e-05 evaluation reward: 3.07
episode: 1258 score: 5.0 memory length: 251042 epsilon: 0.7009348600
064924 steps: 281 lr: 4e-05 evaluation reward: 3.1
```

```
episode: 1259  score: 5.0  memory length: 251335  epsilon: 0.7003547200
06505  steps: 293  lr: 4e-05  evaluation reward: 3.1
episode: 1260  score: 1.0  memory length: 251504  epsilon: 0.7000201000
065123  steps: 169  lr: 4e-05  evaluation reward: 3.08
episode: 1261  score: 3.0  memory length: 251733  epsilon: 0.6995666800
065221  steps: 229  lr: 4e-05  evaluation reward: 3.06
episode: 1262  score: 2.0  memory length: 251931  epsilon: 0.6991746400
065306  steps: 198  lr: 4e-05  evaluation reward: 3.05
episode: 1263  score: 3.0  memory length: 252178  epsilon: 0.6986855800
065412  steps: 247  lr: 4e-05  evaluation reward: 3.07
episode: 1264  score: 0.0  memory length: 252301  epsilon: 0.6984420400
065465  steps: 123  lr: 4e-05  evaluation reward: 3.01
episode: 1265  score: 2.0  memory length: 252519  epsilon: 0.6980104000
065559  steps: 218  lr: 4e-05  evaluation reward: 2.99
episode: 1266  score: 1.0  memory length: 252670  epsilon: 0.6977114200
065624  steps: 151  lr: 4e-05  evaluation reward: 2.95
episode: 1267  score: 7.0  memory length: 253038  epsilon: 0.6969827800
065782  steps: 368  lr: 4e-05  evaluation reward: 2.99
episode: 1268  score: 3.0  memory length: 253269  epsilon: 0.6965254000
065881  steps: 231  lr: 4e-05  evaluation reward: 3.0
episode: 1269  score: 5.0  memory length: 253575  epsilon: 0.6959195200
066013  steps: 306  lr: 4e-05  evaluation reward: 3.02
episode: 1270  score: 4.0  memory length: 253869  epsilon: 0.6953374000
066139  steps: 294  lr: 4e-05  evaluation reward: 3.01
episode: 1271  score: 2.0  memory length: 254050  epsilon: 0.6949790200
066217  steps: 181  lr: 4e-05  evaluation reward: 3.02
episode: 1272  score: 2.0  memory length: 254266  epsilon: 0.6945513400
06631  steps: 216  lr: 4e-05  evaluation reward: 3.02
episode: 1273  score: 4.0  memory length: 254522  epsilon: 0.6940444600
06642  steps: 256  lr: 4e-05  evaluation reward: 3.0
episode: 1274  score: 5.0  memory length: 254856  epsilon: 0.6933831400
066564  steps: 334  lr: 4e-05  evaluation reward: 3.02
episode: 1275  score: 4.0  memory length: 255131  epsilon: 0.6928386400
066682  steps: 275  lr: 4e-05  evaluation reward: 3.02
episode: 1276  score: 2.0  memory length: 255350  epsilon: 0.6924050200
066776  steps: 219  lr: 4e-05  evaluation reward: 3.0
episode: 1277  score: 4.0  memory length: 255626  epsilon: 0.6918585400
066894  steps: 276  lr: 4e-05  evaluation reward: 3.01
episode: 1278  score: 5.0  memory length: 255921  epsilon: 0.6912744400
067021  steps: 295  lr: 4e-05  evaluation reward: 3.02
episode: 1279  score: 2.0  memory length: 256119  epsilon: 0.6908824000
067106  steps: 198  lr: 4e-05  evaluation reward: 3.04
episode: 1280  score: 3.0  memory length: 256364  epsilon: 0.6903973000
067212  steps: 245  lr: 4e-05  evaluation reward: 3.04
episode: 1281  score: 3.0  memory length: 256593  epsilon: 0.6899438800
06731  steps: 229  lr: 4e-05  evaluation reward: 3.06
episode: 1282  score: 2.0  memory length: 256811  epsilon: 0.6895122400
067404  steps: 218  lr: 4e-05  evaluation reward: 3.06
episode: 1283  score: 3.0  memory length: 257057  epsilon: 0.6890251600
06751  steps: 246  lr: 4e-05  evaluation reward: 3.06
episode: 1284  score: 2.0  memory length: 257255  epsilon: 0.6886331200
067595  steps: 198  lr: 4e-05  evaluation reward: 3.05
episode: 1285  score: 6.0  memory length: 257639  epsilon: 0.6878728000
```

```
06776    steps: 384    lr: 4e-05    evaluation reward: 3.11
episode: 1286    score: 3.0    memory length: 257884    epsilon: 0.6873877000
067865    steps: 245    lr: 4e-05    evaluation reward: 3.14
episode: 1287    score: 4.0    memory length: 258161    epsilon: 0.6868392400
067984    steps: 277    lr: 4e-05    evaluation reward: 3.16
episode: 1288    score: 9.0    memory length: 258494    epsilon: 0.6861799000
068127    steps: 333    lr: 4e-05    evaluation reward: 3.24
episode: 1289    score: 4.0    memory length: 258790    epsilon: 0.6855938200
068254    steps: 296    lr: 4e-05    evaluation reward: 3.26
episode: 1290    score: 3.0    memory length: 259018    epsilon: 0.6851423800
068352    steps: 228    lr: 4e-05    evaluation reward: 3.27
episode: 1291    score: 3.0    memory length: 259244    epsilon: 0.6846949000
06845    steps: 226    lr: 4e-05    evaluation reward: 3.28
episode: 1292    score: 4.0    memory length: 259499    epsilon: 0.6841900000
068559    steps: 255    lr: 4e-05    evaluation reward: 3.18
episode: 1293    score: 3.0    memory length: 259745    epsilon: 0.6837029200
068665    steps: 246    lr: 4e-05    evaluation reward: 3.2
episode: 1294    score: 5.0    memory length: 260050    epsilon: 0.6830990200
068796    steps: 305    lr: 4e-05    evaluation reward: 3.22
episode: 1295    score: 4.0    memory length: 260349    epsilon: 0.6825070000
068925    steps: 299    lr: 4e-05    evaluation reward: 3.23
episode: 1296    score: 3.0    memory length: 260574    epsilon: 0.6820615000
069021    steps: 225    lr: 4e-05    evaluation reward: 3.2
episode: 1297    score: 5.0    memory length: 260900    epsilon: 0.6814160200
069161    steps: 326    lr: 4e-05    evaluation reward: 3.21
episode: 1298    score: 4.0    memory length: 261177    epsilon: 0.6808675600
06928    steps: 277    lr: 4e-05    evaluation reward: 3.23
episode: 1299    score: 3.0    memory length: 261403    epsilon: 0.6804200800
069378    steps: 226    lr: 4e-05    evaluation reward: 3.21
episode: 1300    score: 6.0    memory length: 261759    epsilon: 0.6797152000
069531    steps: 356    lr: 4e-05    evaluation reward: 3.27
episode: 1301    score: 3.0    memory length: 261969    epsilon: 0.6792994000
069621    steps: 210    lr: 4e-05    evaluation reward: 3.27
episode: 1302    score: 4.0    memory length: 262247    epsilon: 0.6787489600
06974    steps: 278    lr: 4e-05    evaluation reward: 3.29
episode: 1303    score: 5.0    memory length: 262568    epsilon: 0.6781133800
069878    steps: 321    lr: 4e-05    evaluation reward: 3.3
episode: 1304    score: 3.0    memory length: 262817    epsilon: 0.6776203600
069985    steps: 249    lr: 4e-05    evaluation reward: 3.3
episode: 1305    score: 4.0    memory length: 263071    epsilon: 0.6771174400
070095    steps: 254    lr: 4e-05    evaluation reward: 3.31
episode: 1306    score: 3.0    memory length: 263298    epsilon: 0.6766679800
070192    steps: 227    lr: 4e-05    evaluation reward: 3.33
episode: 1307    score: 3.0    memory length: 263541    epsilon: 0.6761868400
070297    steps: 243    lr: 4e-05    evaluation reward: 3.33
episode: 1308    score: 3.0    memory length: 263768    epsilon: 0.6757373800
070394    steps: 227    lr: 4e-05    evaluation reward: 3.33
episode: 1309    score: 4.0    memory length: 264043    epsilon: 0.6751928800
070512    steps: 275    lr: 4e-05    evaluation reward: 3.34
episode: 1310    score: 6.0    memory length: 264407    epsilon: 0.6744721600
070669    steps: 364    lr: 4e-05    evaluation reward: 3.35
episode: 1311    score: 1.0    memory length: 264557    epsilon: 0.6741751600
070733    steps: 150    lr: 4e-05    evaluation reward: 3.33
```

```
episode: 1312  score: 2.0  memory length: 264755  epsilon: 0.6737831200
070818  steps: 198  lr: 4e-05  evaluation reward: 3.34
episode: 1313  score: 4.0  memory length: 265053  epsilon: 0.6731930800
070947  steps: 298  lr: 4e-05  evaluation reward: 3.35
episode: 1314  score: 3.0  memory length: 265279  epsilon: 0.6727456000
071044  steps: 226  lr: 4e-05  evaluation reward: 3.34
episode: 1315  score: 8.0  memory length: 265678  epsilon: 0.6719555800
071215  steps: 399  lr: 4e-05  evaluation reward: 3.35
episode: 1316  score: 5.0  memory length: 266003  epsilon: 0.6713120800
071355  steps: 325  lr: 4e-05  evaluation reward: 3.36
episode: 1317  score: 3.0  memory length: 266250  epsilon: 0.6708230200
071461  steps: 247  lr: 4e-05  evaluation reward: 3.35
episode: 1318  score: 2.0  memory length: 266448  epsilon: 0.6704309800
071546  steps: 198  lr: 4e-05  evaluation reward: 3.3
episode: 1319  score: 3.0  memory length: 266657  epsilon: 0.6700171600
071636  steps: 209  lr: 4e-05  evaluation reward: 3.3
episode: 1320  score: 6.0  memory length: 267005  epsilon: 0.6693281200
071786  steps: 348  lr: 4e-05  evaluation reward: 3.31
episode: 1321  score: 2.0  memory length: 267223  epsilon: 0.6688964800
071879  steps: 218  lr: 4e-05  evaluation reward: 3.28
episode: 1322  score: 6.0  memory length: 267579  epsilon: 0.6681916000
072032  steps: 356  lr: 4e-05  evaluation reward: 3.34
episode: 1323  score: 3.0  memory length: 267806  epsilon: 0.6677421400
07213  steps: 227  lr: 4e-05  evaluation reward: 3.31
episode: 1324  score: 2.0  memory length: 268025  epsilon: 0.6673085200
072224  steps: 219  lr: 4e-05  evaluation reward: 3.3
episode: 1325  score: 5.0  memory length: 268332  epsilon: 0.6667006600
072356  steps: 307  lr: 4e-05  evaluation reward: 3.35
episode: 1326  score: 6.0  memory length: 268689  epsilon: 0.6659938000
07251  steps: 357  lr: 4e-05  evaluation reward: 3.39
episode: 1327  score: 3.0  memory length: 268915  epsilon: 0.6655463200
072607  steps: 226  lr: 4e-05  evaluation reward: 3.38
episode: 1328  score: 5.0  memory length: 269200  epsilon: 0.6649820200
072729  steps: 285  lr: 4e-05  evaluation reward: 3.39
episode: 1329  score: 3.0  memory length: 269429  epsilon: 0.6645286000
072828  steps: 229  lr: 4e-05  evaluation reward: 3.4
episode: 1330  score: 5.0  memory length: 269769  epsilon: 0.6638554000
072974  steps: 340  lr: 4e-05  evaluation reward: 3.44
episode: 1331  score: 5.0  memory length: 270113  epsilon: 0.6631742800
073122  steps: 344  lr: 4e-05  evaluation reward: 3.46
episode: 1332  score: 4.0  memory length: 270410  epsilon: 0.6625862200
073249  steps: 297  lr: 4e-05  evaluation reward: 3.42
episode: 1333  score: 3.0  memory length: 270674  epsilon: 0.6620635000
073363  steps: 264  lr: 4e-05  evaluation reward: 3.41
episode: 1334  score: 5.0  memory length: 271018  epsilon: 0.6613823800
073511  steps: 344  lr: 4e-05  evaluation reward: 3.45
episode: 1335  score: 6.0  memory length: 271372  epsilon: 0.6606814600
073663  steps: 354  lr: 4e-05  evaluation reward: 3.48
episode: 1336  score: 1.0  memory length: 271523  epsilon: 0.6603824800
073728  steps: 151  lr: 4e-05  evaluation reward: 3.46
episode: 1337  score: 3.0  memory length: 271771  epsilon: 0.6598914400
073834  steps: 248  lr: 4e-05  evaluation reward: 3.44
episode: 1338  score: 3.0  memory length: 272018  epsilon: 0.6594023800
```



```
07394    steps: 247    lr: 4e-05    evaluation reward: 3.41
episode: 1339    score: 5.0    memory length: 272364    epsilon: 0.6587173000
074089    steps: 346    lr: 4e-05    evaluation reward: 3.46
episode: 1340    score: 4.0    memory length: 272604    epsilon: 0.6582421000
074192    steps: 240    lr: 4e-05    evaluation reward: 3.49
episode: 1341    score: 6.0    memory length: 272956    epsilon: 0.6575451400
074344    steps: 352    lr: 4e-05    evaluation reward: 3.53
episode: 1342    score: 5.0    memory length: 273265    epsilon: 0.6569333200
074476    steps: 309    lr: 4e-05    evaluation reward: 3.56
episode: 1343    score: 5.0    memory length: 273569    epsilon: 0.6563314000
074607    steps: 304    lr: 4e-05    evaluation reward: 3.58
episode: 1344    score: 7.0    memory length: 273996    epsilon: 0.6554859400
074791    steps: 427    lr: 4e-05    evaluation reward: 3.62
episode: 1345    score: 4.0    memory length: 274237    epsilon: 0.6550087600
074894    steps: 241    lr: 4e-05    evaluation reward: 3.66
episode: 1346    score: 4.0    memory length: 274532    epsilon: 0.6544246600
075021    steps: 295    lr: 4e-05    evaluation reward: 3.68
episode: 1347    score: 2.0    memory length: 274732    epsilon: 0.6540286600
075107    steps: 200    lr: 4e-05    evaluation reward: 3.69
episode: 1348    score: 7.0    memory length: 275110    epsilon: 0.6532802200
07527    steps: 378    lr: 4e-05    evaluation reward: 3.72
episode: 1349    score: 3.0    memory length: 275336    epsilon: 0.6528327400
075367    steps: 226    lr: 4e-05    evaluation reward: 3.73
episode: 1350    score: 3.0    memory length: 275579    epsilon: 0.6523516000
075471    steps: 243    lr: 4e-05    evaluation reward: 3.72
episode: 1351    score: 5.0    memory length: 275878    epsilon: 0.6517595800
0756    steps: 299    lr: 4e-05    evaluation reward: 3.72
episode: 1352    score: 0.0    memory length: 276000    epsilon: 0.6515180200
075652    steps: 122    lr: 4e-05    evaluation reward: 3.68
episode: 1353    score: 3.0    memory length: 276246    epsilon: 0.6510309400
075758    steps: 246    lr: 4e-05    evaluation reward: 3.67
episode: 1354    score: 3.0    memory length: 276475    epsilon: 0.6505775200
075856    steps: 229    lr: 4e-05    evaluation reward: 3.7
episode: 1355    score: 7.0    memory length: 276864    epsilon: 0.6498073000
076023    steps: 389    lr: 4e-05    evaluation reward: 3.77
episode: 1356    score: 2.0    memory length: 277062    epsilon: 0.6494152600
076109    steps: 198    lr: 4e-05    evaluation reward: 3.76
episode: 1357    score: 3.0    memory length: 277305    epsilon: 0.6489341200
076213    steps: 243    lr: 4e-05    evaluation reward: 3.76
episode: 1358    score: 3.0    memory length: 277515    epsilon: 0.6485183200
076303    steps: 210    lr: 4e-05    evaluation reward: 3.74
episode: 1359    score: 2.0    memory length: 277734    epsilon: 0.6480847000
076397    steps: 219    lr: 4e-05    evaluation reward: 3.71
episode: 1360    score: 6.0    memory length: 278080    epsilon: 0.6473996200
076546    steps: 346    lr: 4e-05    evaluation reward: 3.76
episode: 1361    score: 3.0    memory length: 278309    epsilon: 0.6469462000
076645    steps: 229    lr: 4e-05    evaluation reward: 3.76
episode: 1362    score: 3.0    memory length: 278520    epsilon: 0.6465284200
076735    steps: 211    lr: 4e-05    evaluation reward: 3.77
episode: 1363    score: 2.0    memory length: 278741    epsilon: 0.6460908400
07683    steps: 221    lr: 4e-05    evaluation reward: 3.76
episode: 1364    score: 5.0    memory length: 279032    epsilon: 0.6455146600
076955    steps: 291    lr: 4e-05    evaluation reward: 3.81
```

```
episode: 1365  score: 5.0  memory length: 279336  epsilon: 0.6449127400
077086  steps: 304  lr: 4e-05  evaluation reward: 3.84
episode: 1366  score: 8.0  memory length: 279821  epsilon: 0.6439524400
077294  steps: 485  lr: 4e-05  evaluation reward: 3.91
episode: 1367  score: 5.0  memory length: 280110  epsilon: 0.6433802200
077419  steps: 289  lr: 4e-05  evaluation reward: 3.89
episode: 1368  score: 0.0  memory length: 280233  epsilon: 0.6431366800
077472  steps: 123  lr: 4e-05  evaluation reward: 3.86
episode: 1369  score: 3.0  memory length: 280479  epsilon: 0.6426496000
077577  steps: 246  lr: 4e-05  evaluation reward: 3.84
episode: 1370  score: 5.0  memory length: 280805  epsilon: 0.6420041200
077717  steps: 326  lr: 4e-05  evaluation reward: 3.85
episode: 1371  score: 4.0  memory length: 281064  epsilon: 0.6414913000
077829  steps: 259  lr: 4e-05  evaluation reward: 3.87
episode: 1372  score: 3.0  memory length: 281277  epsilon: 0.6410695600
07792  steps: 213  lr: 4e-05  evaluation reward: 3.88
episode: 1373  score: 4.0  memory length: 281550  epsilon: 0.6405290200
078038  steps: 273  lr: 4e-05  evaluation reward: 3.88
episode: 1374  score: 7.0  memory length: 281966  epsilon: 0.6397053400
078216  steps: 416  lr: 4e-05  evaluation reward: 3.9
episode: 1375  score: 2.0  memory length: 282148  epsilon: 0.6393449800
078295  steps: 182  lr: 4e-05  evaluation reward: 3.88
episode: 1376  score: 2.0  memory length: 282346  epsilon: 0.6389529400
07838  steps: 198  lr: 4e-05  evaluation reward: 3.88
episode: 1377  score: 9.0  memory length: 282865  epsilon: 0.6379253200
078603  steps: 519  lr: 4e-05  evaluation reward: 3.93
episode: 1378  score: 5.0  memory length: 283171  epsilon: 0.6373194400
078734  steps: 306  lr: 4e-05  evaluation reward: 3.93
episode: 1379  score: 2.0  memory length: 283353  epsilon: 0.6369590800
078813  steps: 182  lr: 4e-05  evaluation reward: 3.93
episode: 1380  score: 5.0  memory length: 283642  epsilon: 0.6363868600
078937  steps: 289  lr: 4e-05  evaluation reward: 3.95
episode: 1381  score: 3.0  memory length: 283888  epsilon: 0.6358997800
079043  steps: 246  lr: 4e-05  evaluation reward: 3.95
episode: 1382  score: 12.0  memory length: 284345  epsilon: 0.634994920
0079239  steps: 457  lr: 4e-05  evaluation reward: 4.05
episode: 1383  score: 3.0  memory length: 284595  epsilon: 0.6344999200
079346  steps: 250  lr: 4e-05  evaluation reward: 4.05
episode: 1384  score: 3.0  memory length: 284821  epsilon: 0.6340524400
079444  steps: 226  lr: 4e-05  evaluation reward: 4.06
episode: 1385  score: 4.0  memory length: 285114  epsilon: 0.6334723000
07957  steps: 293  lr: 4e-05  evaluation reward: 4.04
episode: 1386  score: 5.0  memory length: 285450  epsilon: 0.6328070200
079714  steps: 336  lr: 4e-05  evaluation reward: 4.06
episode: 1387  score: 5.0  memory length: 285765  epsilon: 0.6321833200
079849  steps: 315  lr: 4e-05  evaluation reward: 4.07
episode: 1388  score: 4.0  memory length: 286061  epsilon: 0.6315972400
079977  steps: 296  lr: 4e-05  evaluation reward: 4.02
episode: 1389  score: 4.0  memory length: 286301  epsilon: 0.6311220400
08008  steps: 240  lr: 4e-05  evaluation reward: 4.02
episode: 1390  score: 3.0  memory length: 286547  epsilon: 0.6306349600
080186  steps: 246  lr: 4e-05  evaluation reward: 4.02
episode: 1391  score: 1.0  memory length: 286698  epsilon: 0.6303359800
```

```
08025    steps: 151    lr: 4e-05    evaluation reward: 4.0
episode: 1392    score: 1.0    memory length: 286848    epsilon: 0.6300389800
080315    steps: 150    lr: 4e-05    evaluation reward: 3.97
episode: 1393    score: 6.0    memory length: 287211    epsilon: 0.6293202400
080471    steps: 363    lr: 4e-05    evaluation reward: 4.0
episode: 1394    score: 5.0    memory length: 287539    epsilon: 0.6286708000
080612    steps: 328    lr: 4e-05    evaluation reward: 4.0
episode: 1395    score: 3.0    memory length: 287769    epsilon: 0.6282154000
080711    steps: 230    lr: 4e-05    evaluation reward: 3.99
episode: 1396    score: 1.0    memory length: 287920    epsilon: 0.6279164200
080776    steps: 151    lr: 4e-05    evaluation reward: 3.97
episode: 1397    score: 2.0    memory length: 288118    epsilon: 0.6275243800
080861    steps: 198    lr: 4e-05    evaluation reward: 3.94
episode: 1398    score: 0.0    memory length: 288240    epsilon: 0.6272828200
080913    steps: 122    lr: 4e-05    evaluation reward: 3.9
episode: 1399    score: 5.0    memory length: 288563    epsilon: 0.6266432800
081052    steps: 323    lr: 4e-05    evaluation reward: 3.92
episode: 1400    score: 5.0    memory length: 288884    epsilon: 0.6260077000
081119    steps: 321    lr: 4e-05    evaluation reward: 3.91
episode: 1401    score: 3.0    memory length: 289111    epsilon: 0.6255582400
081288    steps: 227    lr: 4e-05    evaluation reward: 3.91
episode: 1402    score: 6.0    memory length: 289471    epsilon: 0.6248454400
081442    steps: 360    lr: 4e-05    evaluation reward: 3.93
episode: 1403    score: 4.0    memory length: 289730    epsilon: 0.6243326200
081554    steps: 259    lr: 4e-05    evaluation reward: 3.92
episode: 1404    score: 3.0    memory length: 289956    epsilon: 0.6238851400
081651    steps: 226    lr: 4e-05    evaluation reward: 3.92
episode: 1405    score: 5.0    memory length: 290300    epsilon: 0.6232040200
081799    steps: 344    lr: 4e-05    evaluation reward: 3.93
episode: 1406    score: 2.0    memory length: 290518    epsilon: 0.6227723800
081892    steps: 218    lr: 4e-05    evaluation reward: 3.92
episode: 1407    score: 2.0    memory length: 290697    epsilon: 0.6224179600
081969    steps: 179    lr: 4e-05    evaluation reward: 3.91
episode: 1408    score: 2.0    memory length: 290913    epsilon: 0.6219902800
082062    steps: 216    lr: 4e-05    evaluation reward: 3.9
episode: 1409    score: 5.0    memory length: 291253    epsilon: 0.6213170800
082208    steps: 340    lr: 4e-05    evaluation reward: 3.91
episode: 1410    score: 4.0    memory length: 291531    epsilon: 0.6207666400
082328    steps: 278    lr: 4e-05    evaluation reward: 3.89
episode: 1411    score: 4.0    memory length: 291806    epsilon: 0.6202221400
082446    steps: 275    lr: 4e-05    evaluation reward: 3.92
episode: 1412    score: 4.0    memory length: 292119    epsilon: 0.6196024000
082581    steps: 313    lr: 4e-05    evaluation reward: 3.94
episode: 1413    score: 6.0    memory length: 292489    epsilon: 0.6188698000
08274    steps: 370    lr: 4e-05    evaluation reward: 3.96
episode: 1414    score: 3.0    memory length: 292735    epsilon: 0.6183827200
082845    steps: 246    lr: 4e-05    evaluation reward: 3.96
episode: 1415    score: 3.0    memory length: 292982    epsilon: 0.6178936600
082952    steps: 247    lr: 4e-05    evaluation reward: 3.91
episode: 1416    score: 2.0    memory length: 293179    epsilon: 0.6175036000
083036    steps: 197    lr: 4e-05    evaluation reward: 3.88
episode: 1417    score: 2.0    memory length: 293377    epsilon: 0.6171115600
083121    steps: 198    lr: 4e-05    evaluation reward: 3.87
```

```
episode: 1418  score: 3.0  memory length: 293609  epsilon: 0.6166522000
083221  steps: 232  lr: 4e-05  evaluation reward: 3.88
episode: 1419  score: 5.0  memory length: 293882  epsilon: 0.6161116600
083338  steps: 273  lr: 4e-05  evaluation reward: 3.9
episode: 1420  score: 4.0  memory length: 294142  epsilon: 0.6155968600
083445  steps: 260  lr: 4e-05  evaluation reward: 3.88
episode: 1421  score: 6.0  memory length: 294497  epsilon: 0.6148939600
083603  steps: 355  lr: 4e-05  evaluation reward: 3.92
episode: 1422  score: 3.0  memory length: 294725  epsilon: 0.6144425200
083701  steps: 228  lr: 4e-05  evaluation reward: 3.89
episode: 1423  score: 5.0  memory length: 295053  epsilon: 0.6137930800
083842  steps: 328  lr: 4e-05  evaluation reward: 3.91
episode: 1424  score: 1.0  memory length: 295224  epsilon: 0.6134545000
083915  steps: 171  lr: 4e-05  evaluation reward: 3.9
episode: 1425  score: 3.0  memory length: 295454  epsilon: 0.6129991000
084014  steps: 230  lr: 4e-05  evaluation reward: 3.88
episode: 1426  score: 6.0  memory length: 295827  epsilon: 0.6122605600
084174  steps: 373  lr: 4e-05  evaluation reward: 3.88
episode: 1427  score: 2.0  memory length: 296045  epsilon: 0.6118289200
084268  steps: 218  lr: 4e-05  evaluation reward: 3.87
episode: 1428  score: 2.0  memory length: 296263  epsilon: 0.6113972800
084362  steps: 218  lr: 4e-05  evaluation reward: 3.84
episode: 1429  score: 3.0  memory length: 296493  epsilon: 0.6109418800
084461  steps: 230  lr: 4e-05  evaluation reward: 3.84
episode: 1430  score: 10.0  memory length: 296994  epsilon: 0.609949900
0084676  steps: 501  lr: 4e-05  evaluation reward: 3.89
episode: 1431  score: 5.0  memory length: 297338  epsilon: 0.6092687800
084824  steps: 344  lr: 4e-05  evaluation reward: 3.89
episode: 1432  score: 3.0  memory length: 297549  epsilon: 0.6088510000
084915  steps: 211  lr: 4e-05  evaluation reward: 3.88
episode: 1433  score: 3.0  memory length: 297778  epsilon: 0.6083975800
085013  steps: 229  lr: 4e-05  evaluation reward: 3.88
episode: 1434  score: 3.0  memory length: 298025  epsilon: 0.6079085200
085119  steps: 247  lr: 4e-05  evaluation reward: 3.86
episode: 1435  score: 5.0  memory length: 298334  epsilon: 0.6072967000
085252  steps: 309  lr: 4e-05  evaluation reward: 3.85
episode: 1436  score: 3.0  memory length: 298545  epsilon: 0.6068789200
085343  steps: 211  lr: 4e-05  evaluation reward: 3.87
episode: 1437  score: 4.0  memory length: 298802  epsilon: 0.6063700600
085453  steps: 257  lr: 4e-05  evaluation reward: 3.88
episode: 1438  score: 6.0  memory length: 299161  epsilon: 0.6056592400
085608  steps: 359  lr: 4e-05  evaluation reward: 3.91
episode: 1439  score: 6.0  memory length: 299479  epsilon: 0.6050296000
085744  steps: 318  lr: 4e-05  evaluation reward: 3.92
episode: 1440  score: 2.0  memory length: 299676  epsilon: 0.6046395400
085829  steps: 197  lr: 4e-05  evaluation reward: 3.9
episode: 1441  score: 7.0  memory length: 300070  epsilon: 0.6038594200
085998  steps: 394  lr: 1.60000000000000003e-05  evaluation reward:
3.91
episode: 1442  score: 4.0  memory length: 300313  epsilon: 0.6033782800
086103  steps: 243  lr: 1.60000000000000003e-05  evaluation reward:
3.9
episode: 1443  score: 4.0  memory length: 300588  epsilon: 0.6028337800
```

086221	steps: 275	lr: 1.6000000000000003e-05	evaluation reward:
3.89			
episode: 1444	score: 6.0	memory length: 300948	epsilon: 0.6021209800
086376	steps: 360	lr: 1.6000000000000003e-05	evaluation reward:
3.88			
episode: 1445	score: 3.0	memory length: 301178	epsilon: 0.6016655800
086474	steps: 230	lr: 1.6000000000000003e-05	evaluation reward:
3.87			
episode: 1446	score: 5.0	memory length: 301484	epsilon: 0.6010597000
086606	steps: 306	lr: 1.6000000000000003e-05	evaluation reward:
3.88			
episode: 1447	score: 2.0	memory length: 301687	epsilon: 0.6006577600
086693	steps: 203	lr: 1.6000000000000003e-05	evaluation reward:
3.88			
episode: 1448	score: 3.0	memory length: 301918	epsilon: 0.6002003800
086793	steps: 231	lr: 1.6000000000000003e-05	evaluation reward:
3.84			
episode: 1449	score: 5.0	memory length: 302246	epsilon: 0.5995509400
086934	steps: 328	lr: 1.6000000000000003e-05	evaluation reward:
3.86			
episode: 1450	score: 4.0	memory length: 302525	epsilon: 0.5989985200
087053	steps: 279	lr: 1.6000000000000003e-05	evaluation reward:
3.87			
episode: 1451	score: 1.0	memory length: 302695	epsilon: 0.5986619200
087127	steps: 170	lr: 1.6000000000000003e-05	evaluation reward:
3.83			
episode: 1452	score: 2.0	memory length: 302913	epsilon: 0.5982302800
08722	steps: 218	lr: 1.6000000000000003e-05	evaluation reward: 3
.85			
episode: 1453	score: 3.0	memory length: 303158	epsilon: 0.5977451800
087326	steps: 245	lr: 1.6000000000000003e-05	evaluation reward:
3.85			
episode: 1454	score: 4.0	memory length: 303435	epsilon: 0.5971967200
087445	steps: 277	lr: 1.6000000000000003e-05	evaluation reward:
3.86			
episode: 1455	score: 3.0	memory length: 303697	epsilon: 0.5966779600
087557	steps: 262	lr: 1.6000000000000003e-05	evaluation reward:
3.82			
episode: 1456	score: 9.0	memory length: 304114	epsilon: 0.5958523000
087736	steps: 417	lr: 1.6000000000000003e-05	evaluation reward:
3.89			
episode: 1457	score: 6.0	memory length: 304509	epsilon: 0.5950702000
087906	steps: 395	lr: 1.6000000000000003e-05	evaluation reward:
3.92			
episode: 1458	score: 3.0	memory length: 304722	epsilon: 0.5946484600
087998	steps: 213	lr: 1.6000000000000003e-05	evaluation reward:
3.92			
episode: 1459	score: 1.0	memory length: 304873	epsilon: 0.5943494800
088063	steps: 151	lr: 1.6000000000000003e-05	evaluation reward:
3.91			
episode: 1460	score: 9.0	memory length: 305339	epsilon: 0.5934268000
088263	steps: 466	lr: 1.6000000000000003e-05	evaluation reward:
3.94			

episode: 1461	score: 9.0	memory length: 305830	epsilon: 0.5924546200
088474	steps: 491	lr: 1.6000000000000003e-05	evaluation reward:
4.0			
episode: 1462	score: 6.0	memory length: 306191	epsilon: 0.5917398400
088629	steps: 361	lr: 1.6000000000000003e-05	evaluation reward:
4.03			
episode: 1463	score: 5.0	memory length: 306498	epsilon: 0.5911319800
088761	steps: 307	lr: 1.6000000000000003e-05	evaluation reward:
4.06			
episode: 1464	score: 9.0	memory length: 306960	epsilon: 0.5902172200
08896	steps: 462	lr: 1.6000000000000003e-05	evaluation reward: 4
.1			
episode: 1465	score: 3.0	memory length: 307171	epsilon: 0.5897994400
08905	steps: 211	lr: 1.6000000000000003e-05	evaluation reward: 4
.08			
episode: 1466	score: 3.0	memory length: 307399	epsilon: 0.5893480000
089149	steps: 228	lr: 1.6000000000000003e-05	evaluation reward:
4.03			
episode: 1467	score: 4.0	memory length: 307677	epsilon: 0.5887975600
089268	steps: 278	lr: 1.6000000000000003e-05	evaluation reward:
4.02			
episode: 1468	score: 5.0	memory length: 307946	epsilon: 0.5882649400
089384	steps: 269	lr: 1.6000000000000003e-05	evaluation reward:
4.07			
episode: 1469	score: 5.0	memory length: 308312	epsilon: 0.5875402600
089541	steps: 366	lr: 1.6000000000000003e-05	evaluation reward:
4.09			
episode: 1470	score: 0.0	memory length: 308435	epsilon: 0.5872967200
089594	steps: 123	lr: 1.6000000000000003e-05	evaluation reward:
4.04			
episode: 1471	score: 4.0	memory length: 308714	epsilon: 0.5867443000
089714	steps: 279	lr: 1.6000000000000003e-05	evaluation reward:
4.04			
episode: 1472	score: 6.0	memory length: 309055	epsilon: 0.5860691200
08986	steps: 341	lr: 1.6000000000000003e-05	evaluation reward: 4
.07			
episode: 1473	score: 4.0	memory length: 309331	epsilon: 0.5855226400
089979	steps: 276	lr: 1.6000000000000003e-05	evaluation reward:
4.07			
episode: 1474	score: 5.0	memory length: 309638	epsilon: 0.5849147800
090111	steps: 307	lr: 1.6000000000000003e-05	evaluation reward:
4.05			
episode: 1475	score: 3.0	memory length: 309869	epsilon: 0.5844574000
09021	steps: 231	lr: 1.6000000000000003e-05	evaluation reward: 4
.06			
episode: 1476	score: 5.0	memory length: 310155	epsilon: 0.5838911200
090333	steps: 286	lr: 1.6000000000000003e-05	evaluation reward:
4.09			
episode: 1477	score: 4.0	memory length: 310429	epsilon: 0.5833486000
090451	steps: 274	lr: 1.6000000000000003e-05	evaluation reward:
4.04			
episode: 1478	score: 6.0	memory length: 310803	epsilon: 0.5826080800
090612	steps: 374	lr: 1.6000000000000003e-05	evaluation reward:

4.05	episode: 1479	score: 4.0	memory length: 311062	epsilon: 0.5820952600
090723	steps: 259	lr: 1.6000000000000003e-05	evaluation reward:	
4.07	episode: 1480	score: 9.0	memory length: 311517	epsilon: 0.5811943600
090919	steps: 455	lr: 1.6000000000000003e-05	evaluation reward:	
4.11	episode: 1481	score: 7.0	memory length: 311921	epsilon: 0.5803944400
091092	steps: 404	lr: 1.6000000000000003e-05	evaluation reward:	
4.15	episode: 1482	score: 4.0	memory length: 312182	epsilon: 0.5798776600
091204	steps: 261	lr: 1.6000000000000003e-05	evaluation reward:	
4.07	episode: 1483	score: 1.0	memory length: 312352	epsilon: 0.5795410600
091278	steps: 170	lr: 1.6000000000000003e-05	evaluation reward:	
4.05	episode: 1484	score: 4.0	memory length: 312613	epsilon: 0.5790242800
09139	steps: 261	lr: 1.6000000000000003e-05	evaluation reward: 4	
.06	episode: 1485	score: 5.0	memory length: 312905	epsilon: 0.5784461200
091515	steps: 292	lr: 1.6000000000000003e-05	evaluation reward:	
4.07	episode: 1486	score: 8.0	memory length: 313308	epsilon: 0.5776481800
091688	steps: 403	lr: 1.6000000000000003e-05	evaluation reward:	
4.1	episode: 1487	score: 8.0	memory length: 313747	epsilon: 0.5767789600
091877	steps: 439	lr: 1.6000000000000003e-05	evaluation reward:	
4.13	episode: 1488	score: 5.0	memory length: 314072	epsilon: 0.5761354600
092017	steps: 325	lr: 1.6000000000000003e-05	evaluation reward:	
4.14	episode: 1489	score: 5.0	memory length: 314395	epsilon: 0.5754959200
092156	steps: 323	lr: 1.6000000000000003e-05	evaluation reward:	
4.15	episode: 1490	score: 6.0	memory length: 314755	epsilon: 0.5747831200
09231	steps: 360	lr: 1.6000000000000003e-05	evaluation reward: 4	
.18	episode: 1491	score: 4.0	memory length: 315049	epsilon: 0.5742010000
092437	steps: 294	lr: 1.6000000000000003e-05	evaluation reward:	
4.21	episode: 1492	score: 2.0	memory length: 315246	epsilon: 0.5738109400
092521	steps: 197	lr: 1.6000000000000003e-05	evaluation reward:	
4.22	episode: 1493	score: 6.0	memory length: 315584	epsilon: 0.5731417000
092667	steps: 338	lr: 1.6000000000000003e-05	evaluation reward:	
4.22	episode: 1494	score: 4.0	memory length: 315843	epsilon: 0.5726288800
092778	steps: 259	lr: 1.6000000000000003e-05	evaluation reward:	
4.21	episode: 1495	score: 3.0	memory length: 316069	epsilon: 0.5721814000
092875	steps: 226	lr: 1.6000000000000003e-05	evaluation reward:	
4.21	episode: 1496	score: 3.0	memory length: 316300	epsilon: 0.5717240200

092975	steps: 231	lr: 1.6000000000000003e-05	evaluation reward:
4.23			
episode: 1497	score: 1.0	memory length: 316451	epsilon: 0.5714250400
093039	steps: 151	lr: 1.6000000000000003e-05	evaluation reward:
4.22			
episode: 1498	score: 3.0	memory length: 316699	epsilon: 0.5709340000
093146	steps: 248	lr: 1.6000000000000003e-05	evaluation reward:
4.25			
episode: 1499	score: 3.0	memory length: 316945	epsilon: 0.5704469200
093252	steps: 246	lr: 1.6000000000000003e-05	evaluation reward:
4.23			
episode: 1500	score: 2.0	memory length: 317146	epsilon: 0.5700489400
093338	steps: 201	lr: 1.6000000000000003e-05	evaluation reward:
4.2			
episode: 1501	score: 3.0	memory length: 317359	epsilon: 0.5696272000
09343	steps: 213	lr: 1.6000000000000003e-05	evaluation reward: 4
.2			
episode: 1502	score: 1.0	memory length: 317510	epsilon: 0.5693282200
093495	steps: 151	lr: 1.6000000000000003e-05	evaluation reward:
4.15			
episode: 1503	score: 3.0	memory length: 317720	epsilon: 0.5689124200
093585	steps: 210	lr: 1.6000000000000003e-05	evaluation reward:
4.14			
episode: 1504	score: 5.0	memory length: 318048	epsilon: 0.5682629800
093726	steps: 328	lr: 1.6000000000000003e-05	evaluation reward:
4.16			
episode: 1505	score: 4.0	memory length: 318324	epsilon: 0.5677165000
093845	steps: 276	lr: 1.6000000000000003e-05	evaluation reward:
4.15			
episode: 1506	score: 6.0	memory length: 318650	epsilon: 0.5670710200
093985	steps: 326	lr: 1.6000000000000003e-05	evaluation reward:
4.19			
episode: 1507	score: 6.0	memory length: 318985	epsilon: 0.5664077200
094129	steps: 335	lr: 1.6000000000000003e-05	evaluation reward:
4.23			
episode: 1508	score: 3.0	memory length: 319231	epsilon: 0.5659206400
094234	steps: 246	lr: 1.6000000000000003e-05	evaluation reward:
4.24			
episode: 1509	score: 3.0	memory length: 319457	epsilon: 0.5654731600
094332	steps: 226	lr: 1.6000000000000003e-05	evaluation reward:
4.22			
episode: 1510	score: 4.0	memory length: 319750	epsilon: 0.5648930200
094457	steps: 293	lr: 1.6000000000000003e-05	evaluation reward:
4.22			
episode: 1511	score: 5.0	memory length: 320074	epsilon: 0.5642515000
094597	steps: 324	lr: 1.6000000000000003e-05	evaluation reward:
4.23			
episode: 1512	score: 3.0	memory length: 320303	epsilon: 0.5637980800
094695	steps: 229	lr: 1.6000000000000003e-05	evaluation reward:
4.22			
episode: 1513	score: 4.0	memory length: 320581	epsilon: 0.5632476400
094815	steps: 278	lr: 1.6000000000000003e-05	evaluation reward:
4.2			

episode: 1514	score: 2.0	memory length: 320762	epsilon: 0.5628892600
094892	steps: 181	lr: 1.6000000000000003e-05	evaluation reward:
4.19			
episode: 1515	score: 6.0	memory length: 321126	epsilon: 0.5621685400
095049	steps: 364	lr: 1.6000000000000003e-05	evaluation reward:
4.22			
episode: 1516	score: 7.0	memory length: 321553	epsilon: 0.5613230800
095232	steps: 427	lr: 1.6000000000000003e-05	evaluation reward:
4.27			
episode: 1517	score: 4.0	memory length: 321846	epsilon: 0.5607429400
095358	steps: 293	lr: 1.6000000000000003e-05	evaluation reward:
4.29			
episode: 1518	score: 3.0	memory length: 322075	epsilon: 0.5602895200
095457	steps: 229	lr: 1.6000000000000003e-05	evaluation reward:
4.29			
episode: 1519	score: 5.0	memory length: 322423	epsilon: 0.5596004800
095606	steps: 348	lr: 1.6000000000000003e-05	evaluation reward:
4.29			
episode: 1520	score: 5.0	memory length: 322747	epsilon: 0.5589589600
095746	steps: 324	lr: 1.6000000000000003e-05	evaluation reward:
4.3			
episode: 1521	score: 3.0	memory length: 322973	epsilon: 0.5585114800
095843	steps: 226	lr: 1.6000000000000003e-05	evaluation reward:
4.27			
episode: 1522	score: 6.0	memory length: 323343	epsilon: 0.5577788800
096002	steps: 370	lr: 1.6000000000000003e-05	evaluation reward:
4.3			
episode: 1523	score: 4.0	memory length: 323602	epsilon: 0.5572660600
096113	steps: 259	lr: 1.6000000000000003e-05	evaluation reward:
4.29			
episode: 1524	score: 6.0	memory length: 323945	epsilon: 0.5565869200
096261	steps: 343	lr: 1.6000000000000003e-05	evaluation reward:
4.34			
episode: 1525	score: 5.0	memory length: 324266	epsilon: 0.5559513400
096399	steps: 321	lr: 1.6000000000000003e-05	evaluation reward:
4.36			
episode: 1526	score: 5.0	memory length: 324613	epsilon: 0.5552642800
096548	steps: 347	lr: 1.6000000000000003e-05	evaluation reward:
4.35			
episode: 1527	score: 6.0	memory length: 324965	epsilon: 0.5545673200
096699	steps: 352	lr: 1.6000000000000003e-05	evaluation reward:
4.39			
episode: 1528	score: 5.0	memory length: 325272	epsilon: 0.5539594600
096831	steps: 307	lr: 1.6000000000000003e-05	evaluation reward:
4.42			
episode: 1529	score: 6.0	memory length: 325627	epsilon: 0.5532565600
096984	steps: 355	lr: 1.6000000000000003e-05	evaluation reward:
4.45			
episode: 1530	score: 3.0	memory length: 325873	epsilon: 0.5527694800
097089	steps: 246	lr: 1.6000000000000003e-05	evaluation reward:
4.38			
episode: 1531	score: 5.0	memory length: 326195	epsilon: 0.5521319200
097228	steps: 322	lr: 1.6000000000000003e-05	evaluation reward:

4.38
episode: 1532 score: 2.0 memory length: 326376 epsilon: 0.5517735400
097306 steps: 181 lr: 1.6000000000000003e-05 evaluation reward:
4.37
episode: 1533 score: 1.0 memory length: 326527 epsilon: 0.5514745600
09737 steps: 151 lr: 1.6000000000000003e-05 evaluation reward: 4
.35
episode: 1534 score: 6.0 memory length: 326880 epsilon: 0.5507756200
097522 steps: 353 lr: 1.6000000000000003e-05 evaluation reward:
4.38
episode: 1535 score: 5.0 memory length: 327184 epsilon: 0.5501737000
097653 steps: 304 lr: 1.6000000000000003e-05 evaluation reward:
4.38
episode: 1536 score: 6.0 memory length: 327574 epsilon: 0.5494015000
09782 steps: 390 lr: 1.6000000000000003e-05 evaluation reward: 4
.41
episode: 1537 score: 5.0 memory length: 327861 epsilon: 0.5488332400
097944 steps: 287 lr: 1.6000000000000003e-05 evaluation reward:
4.42
episode: 1538 score: 5.0 memory length: 328185 epsilon: 0.5481917200
098083 steps: 324 lr: 1.6000000000000003e-05 evaluation reward:
4.41
episode: 1539 score: 5.0 memory length: 328517 epsilon: 0.5475343600
098226 steps: 332 lr: 1.6000000000000003e-05 evaluation reward:
4.4
episode: 1540 score: 4.0 memory length: 328816 epsilon: 0.5469423400
098354 steps: 299 lr: 1.6000000000000003e-05 evaluation reward:
4.42
episode: 1541 score: 7.0 memory length: 329239 epsilon: 0.5461048000
098536 steps: 423 lr: 1.6000000000000003e-05 evaluation reward:
4.42
episode: 1542 score: 2.0 memory length: 329460 epsilon: 0.5456672200
098631 steps: 221 lr: 1.6000000000000003e-05 evaluation reward:
4.4
episode: 1543 score: 4.0 memory length: 329776 epsilon: 0.5450415400
098767 steps: 316 lr: 1.6000000000000003e-05 evaluation reward:
4.4
episode: 1544 score: 5.0 memory length: 330085 epsilon: 0.5444297200
0989 steps: 309 lr: 1.6000000000000003e-05 evaluation reward: 4.
39
episode: 1545 score: 2.0 memory length: 330283 epsilon: 0.5440376800
098985 steps: 198 lr: 1.6000000000000003e-05 evaluation reward:
4.38
episode: 1546 score: 4.0 memory length: 330543 epsilon: 0.5435228800
099097 steps: 260 lr: 1.6000000000000003e-05 evaluation reward:
4.37
episode: 1547 score: 2.0 memory length: 330742 epsilon: 0.5431288600
099182 steps: 199 lr: 1.6000000000000003e-05 evaluation reward:
4.37
episode: 1548 score: 3.0 memory length: 330988 epsilon: 0.5426417800
099288 steps: 246 lr: 1.6000000000000003e-05 evaluation reward:
4.37
episode: 1549 score: 5.0 memory length: 331300 epsilon: 0.5420240200

099422	steps: 312	lr: 1.6000000000000003e-05	evaluation reward:
4.37			
episode: 1550	score: 6.0	memory length: 331650	epsilon: 0.5413310200
099573	steps: 350	lr: 1.6000000000000003e-05	evaluation reward:
4.39			
episode: 1551	score: 5.0	memory length: 331941	epsilon: 0.5407548400
099698	steps: 291	lr: 1.6000000000000003e-05	evaluation reward:
4.43			
episode: 1552	score: 6.0	memory length: 332312	epsilon: 0.5400202600
099857	steps: 371	lr: 1.6000000000000003e-05	evaluation reward:
4.47			
episode: 1553	score: 7.0	memory length: 332736	epsilon: 0.5391807400
100039	steps: 424	lr: 1.6000000000000003e-05	evaluation reward:
4.51			
episode: 1554	score: 6.0	memory length: 333131	epsilon: 0.5383986400
100209	steps: 395	lr: 1.6000000000000003e-05	evaluation reward:
4.53			
episode: 1555	score: 8.0	memory length: 333572	epsilon: 0.5375254600
100399	steps: 441	lr: 1.6000000000000003e-05	evaluation reward:
4.58			
episode: 1556	score: 3.0	memory length: 333803	epsilon: 0.5370680800
100498	steps: 231	lr: 1.6000000000000003e-05	evaluation reward:
4.52			
episode: 1557	score: 5.0	memory length: 334119	epsilon: 0.5364424000
100634	steps: 316	lr: 1.6000000000000003e-05	evaluation reward:
4.51			
episode: 1558	score: 5.0	memory length: 334461	epsilon: 0.5357652400
100781	steps: 342	lr: 1.6000000000000003e-05	evaluation reward:
4.53			
episode: 1559	score: 2.0	memory length: 334661	epsilon: 0.5353692400
100867	steps: 200	lr: 1.6000000000000003e-05	evaluation reward:
4.54			
episode: 1560	score: 5.0	memory length: 334985	epsilon: 0.5347277200
101006	steps: 324	lr: 1.6000000000000003e-05	evaluation reward:
4.5			
episode: 1561	score: 1.0	memory length: 335136	epsilon: 0.5344287400
101071	steps: 151	lr: 1.6000000000000003e-05	evaluation reward:
4.42			
episode: 1562	score: 7.0	memory length: 335546	epsilon: 0.5336169400
101247	steps: 410	lr: 1.6000000000000003e-05	evaluation reward:
4.43			
episode: 1563	score: 8.0	memory length: 335949	epsilon: 0.5328190000
10142	steps: 403	lr: 1.6000000000000003e-05	evaluation reward: 4
.46			
episode: 1564	score: 7.0	memory length: 336353	epsilon: 0.5320190800
101594	steps: 404	lr: 1.6000000000000003e-05	evaluation reward:
4.44			
episode: 1565	score: 8.0	memory length: 336743	epsilon: 0.5312468800
101762	steps: 390	lr: 1.6000000000000003e-05	evaluation reward:
4.49			
episode: 1566	score: 3.0	memory length: 336970	epsilon: 0.5307974200
101859	steps: 227	lr: 1.6000000000000003e-05	evaluation reward:
4.49			

episode: 1567	score: 4.0	memory length: 337227	epsilon: 0.5302885600
10197	steps: 257	lr: 1.6000000000000003e-05	evaluation reward: 4
.49			
episode: 1568	score: 6.0	memory length: 337601	epsilon: 0.5295480400
10213	steps: 374	lr: 1.6000000000000003e-05	evaluation reward: 4
.5			
episode: 1569	score: 5.0	memory length: 337905	epsilon: 0.5289461200
102261	steps: 304	lr: 1.6000000000000003e-05	evaluation reward:
4.5			
episode: 1570	score: 6.0	memory length: 338255	epsilon: 0.5282531200
102412	steps: 350	lr: 1.6000000000000003e-05	evaluation reward:
4.56			
episode: 1571	score: 5.0	memory length: 338577	epsilon: 0.5276155600
10255	steps: 322	lr: 1.6000000000000003e-05	evaluation reward: 4
.57			
episode: 1572	score: 8.0	memory length: 339031	epsilon: 0.5267166400
102745	steps: 454	lr: 1.6000000000000003e-05	evaluation reward:
4.59			
episode: 1573	score: 7.0	memory length: 339450	epsilon: 0.5258870200
102925	steps: 419	lr: 1.6000000000000003e-05	evaluation reward:
4.62			
episode: 1574	score: 8.0	memory length: 339884	epsilon: 0.5250277000
103112	steps: 434	lr: 1.6000000000000003e-05	evaluation reward:
4.65			
episode: 1575	score: 4.0	memory length: 340143	epsilon: 0.5245148800
103223	steps: 259	lr: 1.6000000000000003e-05	evaluation reward:
4.66			
episode: 1576	score: 3.0	memory length: 340369	epsilon: 0.5240674000
10332	steps: 226	lr: 1.6000000000000003e-05	evaluation reward: 4
.64			
episode: 1577	score: 5.0	memory length: 340709	epsilon: 0.5233942000
103466	steps: 340	lr: 1.6000000000000003e-05	evaluation reward:
4.65			
episode: 1578	score: 2.0	memory length: 340907	epsilon: 0.5230021600
103552	steps: 198	lr: 1.6000000000000003e-05	evaluation reward:
4.61			
episode: 1579	score: 8.0	memory length: 341349	epsilon: 0.5221270000
103742	steps: 442	lr: 1.6000000000000003e-05	evaluation reward:
4.65			
episode: 1580	score: 3.0	memory length: 341596	epsilon: 0.5216379400
103848	steps: 247	lr: 1.6000000000000003e-05	evaluation reward:
4.59			
episode: 1581	score: 4.0	memory length: 341876	epsilon: 0.5210835400
103968	steps: 280	lr: 1.6000000000000003e-05	evaluation reward:
4.56			
episode: 1582	score: 3.0	memory length: 342105	epsilon: 0.5206301200
104066	steps: 229	lr: 1.6000000000000003e-05	evaluation reward:
4.55			
episode: 1583	score: 6.0	memory length: 342446	epsilon: 0.5199549400
104213	steps: 341	lr: 1.6000000000000003e-05	evaluation reward:
4.6			
episode: 1584	score: 3.0	memory length: 342712	epsilon: 0.5194282600
104327	steps: 266	lr: 1.6000000000000003e-05	evaluation reward:

```
4.59
episode: 1585    score: 4.0    memory length: 342990    epsilon: 0.5188778200
104447    steps: 278    lr: 1.6000000000000003e-05    evaluation reward:
4.58
episode: 1586    score: 3.0    memory length: 343220    epsilon: 0.5184224200
104546    steps: 230    lr: 1.6000000000000003e-05    evaluation reward:
4.53
episode: 1587    score: 3.0    memory length: 343433    epsilon: 0.5180006800
104637    steps: 213    lr: 1.6000000000000003e-05    evaluation reward:
4.48
episode: 1588    score: 6.0    memory length: 343774    epsilon: 0.5173255000
104784    steps: 341    lr: 1.6000000000000003e-05    evaluation reward:
4.49
episode: 1589    score: 4.0    memory length: 344055    epsilon: 0.5167691200
104905    steps: 281    lr: 1.6000000000000003e-05    evaluation reward:
4.48
episode: 1590    score: 3.0    memory length: 344299    epsilon: 0.5162860000
10501    steps: 244    lr: 1.6000000000000003e-05    evaluation reward: 4
.45
episode: 1591    score: 6.0    memory length: 344693    epsilon: 0.5155058800
105179    steps: 394    lr: 1.6000000000000003e-05    evaluation reward:
4.47
episode: 1592    score: 6.0    memory length: 345113    epsilon: 0.5146742800
10536    steps: 420    lr: 1.6000000000000003e-05    evaluation reward: 4
.51
episode: 1593    score: 4.0    memory length: 345370    epsilon: 0.5141654200
10547    steps: 257    lr: 1.6000000000000003e-05    evaluation reward: 4
.49
episode: 1594    score: 2.0    memory length: 345568    epsilon: 0.5137733800
105555    steps: 198    lr: 1.6000000000000003e-05    evaluation reward:
4.47
episode: 1595    score: 3.0    memory length: 345781    epsilon: 0.5133516400
105647    steps: 213    lr: 1.6000000000000003e-05    evaluation reward:
4.47
episode: 1596    score: 5.0    memory length: 346106    epsilon: 0.5127081400
105786    steps: 325    lr: 1.6000000000000003e-05    evaluation reward:
4.49
episode: 1597    score: 7.0    memory length: 346528    epsilon: 0.5118725800
105968    steps: 422    lr: 1.6000000000000003e-05    evaluation reward:
4.55
episode: 1598    score: 6.0    memory length: 346890    epsilon: 0.5111558200
106123    steps: 362    lr: 1.6000000000000003e-05    evaluation reward:
4.58
episode: 1599    score: 5.0    memory length: 347198    epsilon: 0.5105459800
106256    steps: 308    lr: 1.6000000000000003e-05    evaluation reward:
4.6
episode: 1600    score: 10.0    memory length: 347573    epsilon: 0.509803480
0106417    steps: 375    lr: 1.6000000000000003e-05    evaluation reward:
4.68
episode: 1601    score: 5.0    memory length: 347877    epsilon: 0.5092015600
106548    steps: 304    lr: 1.6000000000000003e-05    evaluation reward:
4.7
episode: 1602    score: 2.0    memory length: 348056    epsilon: 0.5088471400
```

106624	steps: 179	lr: 1.6000000000000003e-05	evaluation reward:
4.71			
episode: 1603	score: 5.0	memory length: 348401	epsilon: 0.5081640400
106773	steps: 345	lr: 1.6000000000000003e-05	evaluation reward:
4.73			
episode: 1604	score: 4.0	memory length: 348680	epsilon: 0.5076116200
106893	steps: 279	lr: 1.6000000000000003e-05	evaluation reward:
4.72			
episode: 1605	score: 1.0	memory length: 348831	epsilon: 0.5073126400
106958	steps: 151	lr: 1.6000000000000003e-05	evaluation reward:
4.69			
episode: 1606	score: 9.0	memory length: 349379	epsilon: 0.5062276000
107193	steps: 548	lr: 1.6000000000000003e-05	evaluation reward:
4.72			
episode: 1607	score: 9.0	memory length: 349900	epsilon: 0.5051960200
107417	steps: 521	lr: 1.6000000000000003e-05	evaluation reward:
4.75			
episode: 1608	score: 3.0	memory length: 350127	epsilon: 0.5047465600
107515	steps: 227	lr: 1.6000000000000003e-05	evaluation reward:
4.75			
episode: 1609	score: 5.0	memory length: 350445	epsilon: 0.5041169200
107651	steps: 318	lr: 1.6000000000000003e-05	evaluation reward:
4.77			
episode: 1610	score: 7.0	memory length: 350852	epsilon: 0.5033110600
107826	steps: 407	lr: 1.6000000000000003e-05	evaluation reward:
4.8			
episode: 1611	score: 1.0	memory length: 351021	epsilon: 0.5029764400
107899	steps: 169	lr: 1.6000000000000003e-05	evaluation reward:
4.76			
episode: 1612	score: 4.0	memory length: 351321	epsilon: 0.5023824400
108028	steps: 300	lr: 1.6000000000000003e-05	evaluation reward:
4.77			
episode: 1613	score: 4.0	memory length: 351599	epsilon: 0.5018320000
108147	steps: 278	lr: 1.6000000000000003e-05	evaluation reward:
4.77			
episode: 1614	score: 7.0	memory length: 351973	epsilon: 0.5010914800
108308	steps: 374	lr: 1.6000000000000003e-05	evaluation reward:
4.82			
episode: 1615	score: 3.0	memory length: 352186	epsilon: 0.5006697400
1084	steps: 213	lr: 1.6000000000000003e-05	evaluation reward: 4.
79			
episode: 1616	score: 5.0	memory length: 352528	epsilon: 0.4999925800
1085445	steps: 342	lr: 1.6000000000000003e-05	evaluation reward:
4.77			
episode: 1617	score: 5.0	memory length: 352849	epsilon: 0.4993570000
108504	steps: 321	lr: 1.6000000000000003e-05	evaluation reward:
4.78			
episode: 1618	score: 5.0	memory length: 353173	epsilon: 0.4987154800
1084637	steps: 324	lr: 1.6000000000000003e-05	evaluation reward:
4.8			
episode: 1619	score: 5.0	memory length: 353496	epsilon: 0.4980759400
108423	steps: 323	lr: 1.6000000000000003e-05	evaluation reward:
4.8			

episode: 1620	score: 4.0	memory length: 353778	epsilon: 0.4975175800
108388	steps: 282	lr: 1.6000000000000003e-05	evaluation reward:
4.79			
episode: 1621	score: 3.0	memory length: 354023	epsilon: 0.4970324800
108357	steps: 245	lr: 1.6000000000000003e-05	evaluation reward:
4.79			
episode: 1622	score: 5.0	memory length: 354327	epsilon: 0.4964305600
108319	steps: 304	lr: 1.6000000000000003e-05	evaluation reward:
4.78			
episode: 1623	score: 6.0	memory length: 354662	epsilon: 0.4957672600
108277	steps: 335	lr: 1.6000000000000003e-05	evaluation reward:
4.8			
episode: 1624	score: 2.0	memory length: 354860	epsilon: 0.4953752200
1082523	steps: 198	lr: 1.6000000000000003e-05	evaluation reward:
4.76			
episode: 1625	score: 7.0	memory length: 355269	epsilon: 0.4945654000
108201	steps: 409	lr: 1.6000000000000003e-05	evaluation reward:
4.78			
episode: 1626	score: 2.0	memory length: 355466	epsilon: 0.4941753400
1081764	steps: 197	lr: 1.6000000000000003e-05	evaluation reward:
4.75			
episode: 1627	score: 4.0	memory length: 355741	epsilon: 0.4936308400
108142	steps: 275	lr: 1.6000000000000003e-05	evaluation reward:
4.73			
episode: 1628	score: 5.0	memory length: 356069	epsilon: 0.4929814000
108101	steps: 328	lr: 1.6000000000000003e-05	evaluation reward:
4.73			
episode: 1629	score: 5.0	memory length: 356393	epsilon: 0.4923398800
1080603	steps: 324	lr: 1.6000000000000003e-05	evaluation reward:
4.72			
episode: 1630	score: 4.0	memory length: 356650	epsilon: 0.4918310200
108028	steps: 257	lr: 1.6000000000000003e-05	evaluation reward:
4.73			
episode: 1631	score: 5.0	memory length: 356990	epsilon: 0.4911578200
1079855	steps: 340	lr: 1.6000000000000003e-05	evaluation reward:
4.73			
episode: 1632	score: 4.0	memory length: 357309	epsilon: 0.4905262000
1079456	steps: 319	lr: 1.6000000000000003e-05	evaluation reward:
4.75			
episode: 1633	score: 6.0	memory length: 357685	epsilon: 0.4897817200
1078984	steps: 376	lr: 1.6000000000000003e-05	evaluation reward:
4.8			
episode: 1634	score: 3.0	memory length: 357897	epsilon: 0.4893619600
107872	steps: 212	lr: 1.6000000000000003e-05	evaluation reward:
4.77			
episode: 1635	score: 5.0	memory length: 358193	epsilon: 0.4887758800
107835	steps: 296	lr: 1.6000000000000003e-05	evaluation reward:
4.77			
episode: 1636	score: 5.0	memory length: 358517	epsilon: 0.4881343600
107794	steps: 324	lr: 1.6000000000000003e-05	evaluation reward:
4.76			
episode: 1637	score: 6.0	memory length: 358862	epsilon: 0.4874512600
107751	steps: 345	lr: 1.6000000000000003e-05	evaluation reward:

4.77
episode: 1638 score: 7.0 memory length: 359287 epsilon: 0.4866097600
107698 steps: 425 lr: 1.6000000000000003e-05 evaluation reward:
4.79
episode: 1639 score: 9.0 memory length: 359769 epsilon: 0.4856554000
1076374 steps: 482 lr: 1.6000000000000003e-05 evaluation reward:
4.83
episode: 1640 score: 4.0 memory length: 360068 epsilon: 0.4850633800
1076 steps: 299 lr: 1.6000000000000003e-05 evaluation reward: 4.
83
episode: 1641 score: 0.0 memory length: 360191 epsilon: 0.4848198400
1075845 steps: 123 lr: 1.6000000000000003e-05 evaluation reward:
4.76
episode: 1642 score: 6.0 memory length: 360526 epsilon: 0.4841565400
1075425 steps: 335 lr: 1.6000000000000003e-05 evaluation reward:
4.8
episode: 1643 score: 3.0 memory length: 360754 epsilon: 0.4837051000
107514 steps: 228 lr: 1.6000000000000003e-05 evaluation reward:
4.79
episode: 1644 score: 9.0 memory length: 361204 epsilon: 0.4828141000
1074576 steps: 450 lr: 1.6000000000000003e-05 evaluation reward:
4.83
episode: 1645 score: 3.0 memory length: 361432 epsilon: 0.4823626600
107429 steps: 228 lr: 1.6000000000000003e-05 evaluation reward:
4.84
episode: 1646 score: 4.0 memory length: 361707 epsilon: 0.4818181600
1073946 steps: 275 lr: 1.6000000000000003e-05 evaluation reward:
4.84
episode: 1647 score: 9.0 memory length: 362163 epsilon: 0.4809152800
1073375 steps: 456 lr: 1.6000000000000003e-05 evaluation reward:
4.91
episode: 1648 score: 5.0 memory length: 362510 epsilon: 0.4802282200
107294 steps: 347 lr: 1.6000000000000003e-05 evaluation reward:
4.93
episode: 1649 score: 3.0 memory length: 362740 epsilon: 0.4797728200
107265 steps: 230 lr: 1.6000000000000003e-05 evaluation reward:
4.91
episode: 1650 score: 14.0 memory length: 363185 epsilon: 0.478891720
01072094 steps: 445 lr: 1.6000000000000003e-05 evaluation reward
: 4.99
episode: 1651 score: 5.0 memory length: 363527 epsilon: 0.4782145600
1071666 steps: 342 lr: 1.6000000000000003e-05 evaluation reward:
4.99
episode: 1652 score: 4.0 memory length: 363802 epsilon: 0.4776700600
107132 steps: 275 lr: 1.6000000000000003e-05 evaluation reward:
4.97
episode: 1653 score: 11.0 memory length: 364365 epsilon: 0.476555320
01070616 steps: 563 lr: 1.6000000000000003e-05 evaluation reward
: 5.01
episode: 1654 score: 5.0 memory length: 364732 epsilon: 0.4758286600
1070157 steps: 367 lr: 1.6000000000000003e-05 evaluation reward:
5.0
episode: 1655 score: 7.0 memory length: 365155 epsilon: 0.4749911200

1069627	steps: 423	lr: 1.6000000000000003e-05	evaluation reward: 4.99
episode: 1656	score: 5.0	memory length: 365498	epsilon: 0.4743119800
1069197	steps: 343	lr: 1.6000000000000003e-05	evaluation reward: 5.01
episode: 1657	score: 8.0	memory length: 365964	epsilon: 0.4733893000
1068613	steps: 466	lr: 1.6000000000000003e-05	evaluation reward: 5.04
episode: 1658	score: 7.0	memory length: 366390	epsilon: 0.4725458200
106808	steps: 426	lr: 1.6000000000000003e-05	evaluation reward: 5.06
episode: 1659	score: 4.0	memory length: 366665	epsilon: 0.4720013200
1067735	steps: 275	lr: 1.6000000000000003e-05	evaluation reward: 5.08
episode: 1660	score: 5.0	memory length: 366962	epsilon: 0.4714132600
1067363	steps: 297	lr: 1.6000000000000003e-05	evaluation reward: 5.08
episode: 1661	score: 6.0	memory length: 367301	epsilon: 0.4707420400
106694	steps: 339	lr: 1.6000000000000003e-05	evaluation reward: 5.13
episode: 1662	score: 5.0	memory length: 367649	epsilon: 0.4700530000
10665	steps: 348	lr: 1.6000000000000003e-05	evaluation reward: 5.11
episode: 1663	score: 5.0	memory length: 367974	epsilon: 0.4694095000
1066095	steps: 325	lr: 1.6000000000000003e-05	evaluation reward: 5.08
episode: 1664	score: 3.0	memory length: 368202	epsilon: 0.4689580600
106581	steps: 228	lr: 1.6000000000000003e-05	evaluation reward: 5.04
episode: 1665	score: 3.0	memory length: 368447	epsilon: 0.4684729600
10655	steps: 245	lr: 1.6000000000000003e-05	evaluation reward: 4.99
episode: 1666	score: 3.0	memory length: 368676	epsilon: 0.4680195400
1065216	steps: 229	lr: 1.6000000000000003e-05	evaluation reward: 4.99
episode: 1667	score: 6.0	memory length: 369040	epsilon: 0.4672988200
106476	steps: 364	lr: 1.6000000000000003e-05	evaluation reward: 5.01
episode: 1668	score: 5.0	memory length: 369327	epsilon: 0.4667305600
10644	steps: 287	lr: 1.6000000000000003e-05	evaluation reward: 5.0
episode: 1669	score: 5.0	memory length: 369669	epsilon: 0.4660534000
106397	steps: 342	lr: 1.6000000000000003e-05	evaluation reward: 5.0
episode: 1670	score: 4.0	memory length: 369963	epsilon: 0.4654712800
1063603	steps: 294	lr: 1.6000000000000003e-05	evaluation reward: 4.98
episode: 1671	score: 11.0	memory length: 370385	epsilon: 0.464635720
01063075	steps: 422	lr: 1.6000000000000003e-05	evaluation reward: 5.04
episode: 1672	score: 4.0	memory length: 370646	epsilon: 0.4641189400
106275	steps: 261	lr: 1.6000000000000003e-05	evaluation reward: 5.0

episode: 1673	score: 5.0	memory length: 370937	epsilon: 0.4635427600
1062383	steps: 291	lr: 1.6000000000000003e-05	evaluation reward: 4.98
episode: 1674	score: 4.0	memory length: 371214	epsilon: 0.4629943000
1062036	steps: 277	lr: 1.6000000000000003e-05	evaluation reward: 4.94
episode: 1675	score: 1.0	memory length: 371365	epsilon: 0.4626953200
1061847	steps: 151	lr: 1.6000000000000003e-05	evaluation reward: 4.91
episode: 1676	score: 9.0	memory length: 371794	epsilon: 0.4618459000
106131	steps: 429	lr: 1.6000000000000003e-05	evaluation reward: 4.97
episode: 1677	score: 1.0	memory length: 371944	epsilon: 0.4615489000
106112	steps: 150	lr: 1.6000000000000003e-05	evaluation reward: 4.93
episode: 1678	score: 11.0	memory length: 372503	epsilon: 0.460442080
0106042	steps: 559	lr: 1.6000000000000003e-05	evaluation reward: 5.02
episode: 1679	score: 4.0	memory length: 372761	epsilon: 0.4599312400
10601	steps: 258	lr: 1.6000000000000003e-05	evaluation reward: 4.98
episode: 1680	score: 5.0	memory length: 373069	epsilon: 0.4593214000
105971	steps: 308	lr: 1.6000000000000003e-05	evaluation reward: 5.0
episode: 1681	score: 6.0	memory length: 373429	epsilon: 0.4586086000
105926	steps: 360	lr: 1.6000000000000003e-05	evaluation reward: 5.02
episode: 1682	score: 10.0	memory length: 373925	epsilon: 0.457626520
0105864	steps: 496	lr: 1.6000000000000003e-05	evaluation reward: 5.09
episode: 1683	score: 3.0	memory length: 374137	epsilon: 0.4572067600
1058375	steps: 212	lr: 1.6000000000000003e-05	evaluation reward: 5.06
episode: 1684	score: 3.0	memory length: 374364	epsilon: 0.4567573000
105809	steps: 227	lr: 1.6000000000000003e-05	evaluation reward: 5.06
episode: 1685	score: 5.0	memory length: 374704	epsilon: 0.4560841000
1057664	steps: 340	lr: 1.6000000000000003e-05	evaluation reward: 5.07
episode: 1686	score: 6.0	memory length: 375068	epsilon: 0.4553633800
105721	steps: 364	lr: 1.6000000000000003e-05	evaluation reward: 5.1
episode: 1687	score: 8.0	memory length: 375482	epsilon: 0.4545436600
105669	steps: 414	lr: 1.6000000000000003e-05	evaluation reward: 5.15
episode: 1688	score: 8.0	memory length: 375932	epsilon: 0.4536526600
1056126	steps: 450	lr: 1.6000000000000003e-05	evaluation reward: 5.17
episode: 1689	score: 3.0	memory length: 376158	epsilon: 0.4532051800
1055843	steps: 226	lr: 1.6000000000000003e-05	evaluation reward: 5.16
episode: 1690	score: 7.0	memory length: 376532	epsilon: 0.4524646600
1055374	steps: 374	lr: 1.6000000000000003e-05	evaluation reward:

5.2
episode: 1691 score: 4.0 memory length: 376811 epsilon: 0.4519122400
1055025 steps: 279 lr: 1.6000000000000003e-05 evaluation reward:
5.18
episode: 1692 score: 5.0 memory length: 377117 epsilon: 0.4513063600
105464 steps: 306 lr: 1.6000000000000003e-05 evaluation reward:
5.17
episode: 1693 score: 3.0 memory length: 377364 epsilon: 0.4508173000
105433 steps: 247 lr: 1.6000000000000003e-05 evaluation reward:
5.16
episode: 1694 score: 2.0 memory length: 377546 epsilon: 0.4504569400
1054104 steps: 182 lr: 1.6000000000000003e-05 evaluation reward:
5.16
episode: 1695 score: 4.0 memory length: 377848 epsilon: 0.4498589800
1053726 steps: 302 lr: 1.6000000000000003e-05 evaluation reward:
5.17
episode: 1696 score: 9.0 memory length: 378301 epsilon: 0.4489620400
105316 steps: 453 lr: 1.6000000000000003e-05 evaluation reward:
5.21
episode: 1697 score: 3.0 memory length: 378534 epsilon: 0.4485007000
1052866 steps: 233 lr: 1.6000000000000003e-05 evaluation reward:
5.17
episode: 1698 score: 3.0 memory length: 378762 epsilon: 0.4480492600
105258 steps: 228 lr: 1.6000000000000003e-05 evaluation reward:
5.14
episode: 1699 score: 7.0 memory length: 379117 epsilon: 0.4473463600
1052136 steps: 355 lr: 1.6000000000000003e-05 evaluation reward:
5.16
episode: 1700 score: 7.0 memory length: 379483 epsilon: 0.4466216800
105168 steps: 366 lr: 1.6000000000000003e-05 evaluation reward:
5.13
episode: 1701 score: 4.0 memory length: 379758 epsilon: 0.4460771800
1051333 steps: 275 lr: 1.6000000000000003e-05 evaluation reward:
5.12
episode: 1702 score: 2.0 memory length: 379937 epsilon: 0.4457227600
105111 steps: 179 lr: 1.6000000000000003e-05 evaluation reward:
5.12
episode: 1703 score: 5.0 memory length: 380222 epsilon: 0.4451584600
105075 steps: 285 lr: 1.6000000000000003e-05 evaluation reward:
5.12
episode: 1704 score: 4.0 memory length: 380481 epsilon: 0.4446456400
105043 steps: 259 lr: 1.6000000000000003e-05 evaluation reward:
5.12
episode: 1705 score: 8.0 memory length: 380884 epsilon: 0.4438477000
104992 steps: 403 lr: 1.6000000000000003e-05 evaluation reward:
5.19
episode: 1706 score: 7.0 memory length: 381308 epsilon: 0.4430081800
104939 steps: 424 lr: 1.6000000000000003e-05 evaluation reward:
5.17
episode: 1707 score: 6.0 memory length: 381658 epsilon: 0.4423151800
1048953 steps: 350 lr: 1.6000000000000003e-05 evaluation reward:
5.14
episode: 1708 score: 3.0 memory length: 381887 epsilon: 0.4418617600

1048666	steps: 229	lr: 1.6000000000000003e-05	evaluation reward:
5.14			
episode: 1709	score: 4.0	memory length: 382162	epsilon: 0.4413172600
104832	steps: 275	lr: 1.6000000000000003e-05	evaluation reward:
5.13			
episode: 1710	score: 9.0	memory length: 382669	epsilon: 0.4403134000
1047686	steps: 507	lr: 1.6000000000000003e-05	evaluation reward:
5.15			
episode: 1711	score: 2.0	memory length: 382888	epsilon: 0.4398797800
104741	steps: 219	lr: 1.6000000000000003e-05	evaluation reward:
5.16			
episode: 1712	score: 5.0	memory length: 383173	epsilon: 0.4393154800
1047055	steps: 285	lr: 1.6000000000000003e-05	evaluation reward:
5.17			
episode: 1713	score: 4.0	memory length: 383451	epsilon: 0.4387650400
1046707	steps: 278	lr: 1.6000000000000003e-05	evaluation reward:
5.17			
episode: 1714	score: 1.0	memory length: 383621	epsilon: 0.4384284400
1046494	steps: 170	lr: 1.6000000000000003e-05	evaluation reward:
5.11			
episode: 1715	score: 9.0	memory length: 384079	epsilon: 0.4375216000
104592	steps: 458	lr: 1.6000000000000003e-05	evaluation reward:
5.17			
episode: 1716	score: 7.0	memory length: 384467	epsilon: 0.4367533600
1045434	steps: 388	lr: 1.6000000000000003e-05	evaluation reward:
5.19			
episode: 1717	score: 3.0	memory length: 384732	epsilon: 0.4362286600
10451	steps: 265	lr: 1.6000000000000003e-05	evaluation reward: 5
.17			
episode: 1718	score: 4.0	memory length: 385026	epsilon: 0.4356465400
1044734	steps: 294	lr: 1.6000000000000003e-05	evaluation reward:
5.16			
episode: 1719	score: 10.0	memory length: 385455	epsilon: 0.434797120
01044196	steps: 429	lr: 1.6000000000000003e-05	evaluation reward
: 5.21			
episode: 1720	score: 4.0	memory length: 385750	epsilon: 0.4342130200
1043827	steps: 295	lr: 1.6000000000000003e-05	evaluation reward:
5.21			
episode: 1721	score: 1.0	memory length: 385919	epsilon: 0.4338784000
1043615	steps: 169	lr: 1.6000000000000003e-05	evaluation reward:
5.19			
episode: 1722	score: 7.0	memory length: 386305	epsilon: 0.4331141200
104313	steps: 386	lr: 1.6000000000000003e-05	evaluation reward:
5.21			
episode: 1723	score: 8.0	memory length: 386729	epsilon: 0.4322746000
10426	steps: 424	lr: 1.6000000000000003e-05	evaluation reward: 5
.23			
episode: 1724	score: 7.0	memory length: 387115	epsilon: 0.4315103200
1042117	steps: 386	lr: 1.6000000000000003e-05	evaluation reward:
5.28			
episode: 1725	score: 3.0	memory length: 387325	epsilon: 0.4310945200
1041854	steps: 210	lr: 1.6000000000000003e-05	evaluation reward:
5.24			

episode: 1726	score: 4.0	memory length: 387564	epsilon: 0.4306213000
1041554	steps: 239	lr: 1.6000000000000003e-05	evaluation reward:
5.26			
episode: 1727	score: 6.0	memory length: 387919	epsilon: 0.4299184000
104111	steps: 355	lr: 1.6000000000000003e-05	evaluation reward:
5.28			
episode: 1728	score: 3.0	memory length: 388132	epsilon: 0.4294966600
104084	steps: 213	lr: 1.6000000000000003e-05	evaluation reward:
5.26			
episode: 1729	score: 8.0	memory length: 388603	epsilon: 0.4285640800
104025	steps: 471	lr: 1.6000000000000003e-05	evaluation reward:
5.29			
episode: 1730	score: 7.0	memory length: 389013	epsilon: 0.4277522800
103974	steps: 410	lr: 1.6000000000000003e-05	evaluation reward:
5.32			
episode: 1731	score: 7.0	memory length: 389404	epsilon: 0.4269781000
103925	steps: 391	lr: 1.6000000000000003e-05	evaluation reward:
5.34			
episode: 1732	score: 9.0	memory length: 389731	epsilon: 0.4263306400
103884	steps: 327	lr: 1.6000000000000003e-05	evaluation reward:
5.39			
episode: 1733	score: 4.0	memory length: 389991	epsilon: 0.4258158400
1038514	steps: 260	lr: 1.6000000000000003e-05	evaluation reward:
5.37			
episode: 1734	score: 7.0	memory length: 390354	epsilon: 0.4250971000
103806	steps: 363	lr: 1.6000000000000003e-05	evaluation reward:
5.41			
episode: 1735	score: 3.0	memory length: 390564	epsilon: 0.4246813000
1037796	steps: 210	lr: 1.6000000000000003e-05	evaluation reward:
5.39			
episode: 1736	score: 4.0	memory length: 390825	epsilon: 0.4241645200
103747	steps: 261	lr: 1.6000000000000003e-05	evaluation reward:
5.38			
episode: 1737	score: 4.0	memory length: 391100	epsilon: 0.4236200200
1037125	steps: 275	lr: 1.6000000000000003e-05	evaluation reward:
5.36			
episode: 1738	score: 4.0	memory length: 391358	epsilon: 0.4231091800
10368	steps: 258	lr: 1.6000000000000003e-05	evaluation reward: 5
.33			
episode: 1739	score: 5.0	memory length: 391663	epsilon: 0.4225052800
103642	steps: 305	lr: 1.6000000000000003e-05	evaluation reward:
5.29			
episode: 1740	score: 3.0	memory length: 391892	epsilon: 0.4220518600
103613	steps: 229	lr: 1.6000000000000003e-05	evaluation reward:
5.28			
episode: 1741	score: 10.0	memory length: 392435	epsilon: 0.420976720
0103545	steps: 543	lr: 1.6000000000000003e-05	evaluation reward:
5.38			
episode: 1742	score: 7.0	memory length: 392801	epsilon: 0.4202520400
1034994	steps: 366	lr: 1.6000000000000003e-05	evaluation reward:
5.39			
episode: 1743	score: 7.0	memory length: 393238	epsilon: 0.4193867800
1034446	steps: 437	lr: 1.6000000000000003e-05	evaluation reward:

5.43
episode: 1744 score: 3.0 memory length: 393505 epsilon: 0.4188581200
103411 steps: 267 lr: 1.6000000000000003e-05 evaluation reward:
5.37
episode: 1745 score: 7.0 memory length: 393924 epsilon: 0.4180285000
1033587 steps: 419 lr: 1.6000000000000003e-05 evaluation reward:
5.41
episode: 1746 score: 5.0 memory length: 394249 epsilon: 0.4173850000
103318 steps: 325 lr: 1.6000000000000003e-05 evaluation reward:
5.42
episode: 1747 score: 3.0 memory length: 394459 epsilon: 0.4169692000
1032917 steps: 210 lr: 1.6000000000000003e-05 evaluation reward:
5.36
episode: 1748 score: 13.0 memory length: 395017 epsilon: 0.415864360
0103222 steps: 558 lr: 1.6000000000000003e-05 evaluation reward:
5.44
episode: 1749 score: 8.0 memory length: 395475 epsilon: 0.4149575200
1031644 steps: 458 lr: 1.6000000000000003e-05 evaluation reward:
5.49
episode: 1750 score: 3.0 memory length: 395704 epsilon: 0.4145041000
1031357 steps: 229 lr: 1.6000000000000003e-05 evaluation reward:
5.38
episode: 1751 score: 13.0 memory length: 396246 epsilon: 0.413430940
0103068 steps: 542 lr: 1.6000000000000003e-05 evaluation reward:
5.46
episode: 1752 score: 6.0 memory length: 396586 epsilon: 0.4127577400
103025 steps: 340 lr: 1.6000000000000003e-05 evaluation reward:
5.48
episode: 1753 score: 4.0 memory length: 396848 epsilon: 0.4122389800
1029924 steps: 262 lr: 1.6000000000000003e-05 evaluation reward:
5.41
episode: 1754 score: 5.0 memory length: 397159 epsilon: 0.4116232000
1029534 steps: 311 lr: 1.6000000000000003e-05 evaluation reward:
5.41
episode: 1755 score: 4.0 memory length: 397433 epsilon: 0.4110806800
102919 steps: 274 lr: 1.6000000000000003e-05 evaluation reward:
5.38
episode: 1756 score: 7.0 memory length: 397820 epsilon: 0.4103144200
1028706 steps: 387 lr: 1.6000000000000003e-05 evaluation reward:
5.4
episode: 1757 score: 7.0 memory length: 398244 epsilon: 0.4094749000
1028175 steps: 424 lr: 1.6000000000000003e-05 evaluation reward:
5.39
episode: 1758 score: 3.0 memory length: 398492 epsilon: 0.4089838600
1027864 steps: 248 lr: 1.6000000000000003e-05 evaluation reward:
5.35
episode: 1759 score: 4.0 memory length: 398769 epsilon: 0.4084354000
102752 steps: 277 lr: 1.6000000000000003e-05 evaluation reward:
5.35
episode: 1760 score: 7.0 memory length: 399192 epsilon: 0.4075978600
102699 steps: 423 lr: 1.6000000000000003e-05 evaluation reward:
5.37
episode: 1761 score: 4.0 memory length: 399443 epsilon: 0.4071008800

1026673	steps: 251	lr: 1.6000000000000003e-05	evaluation reward: 5.35
episode: 1762	score: 5.0	memory length: 399768	epsilon: 0.4064573800
1026266	steps: 325	lr: 1.6000000000000003e-05	evaluation reward: 5.35
episode: 1763	score: 6.0	memory length: 400091	epsilon: 0.4058178400
102586	steps: 323	lr: 6.400000000000001e-06	evaluation reward: 5.36
episode: 1764	score: 6.0	memory length: 400463	epsilon: 0.4050812800
1025395	steps: 372	lr: 6.400000000000001e-06	evaluation reward: 5.39
episode: 1765	score: 6.0	memory length: 400799	epsilon: 0.4044160000
1024974	steps: 336	lr: 6.400000000000001e-06	evaluation reward: 5.42
episode: 1766	score: 7.0	memory length: 401160	epsilon: 0.4037012200
102452	steps: 361	lr: 6.400000000000001e-06	evaluation reward: 5.46
episode: 1767	score: 5.0	memory length: 401438	epsilon: 0.4031507800
1024174	steps: 278	lr: 6.400000000000001e-06	evaluation reward: 5.45
episode: 1768	score: 4.0	memory length: 401713	epsilon: 0.4026062800
102383	steps: 275	lr: 6.400000000000001e-06	evaluation reward: 5.44
episode: 1769	score: 4.0	memory length: 402009	epsilon: 0.4020202000
102346	steps: 296	lr: 6.400000000000001e-06	evaluation reward: 5.43
episode: 1770	score: 6.0	memory length: 402399	epsilon: 0.4012480000
102297	steps: 390	lr: 6.400000000000001e-06	evaluation reward: 5.45
episode: 1771	score: 6.0	memory length: 402705	epsilon: 0.4006421200
1022587	steps: 306	lr: 6.400000000000001e-06	evaluation reward: 5.4
episode: 1772	score: 10.0	memory length: 403215	epsilon: 0.399632320
0102195	steps: 510	lr: 6.400000000000001e-06	evaluation reward: 5.46
episode: 1773	score: 4.0	memory length: 403457	epsilon: 0.3991531600
1021645	steps: 242	lr: 6.400000000000001e-06	evaluation reward: 5.45
episode: 1774	score: 3.0	memory length: 403670	epsilon: 0.3987314200
102138	steps: 213	lr: 6.400000000000001e-06	evaluation reward: 5.44
episode: 1775	score: 6.0	memory length: 404042	epsilon: 0.3979948600
102091	steps: 372	lr: 6.400000000000001e-06	evaluation reward: 5.49
episode: 1776	score: 9.0	memory length: 404534	epsilon: 0.3970207000
1020295	steps: 492	lr: 6.400000000000001e-06	evaluation reward: 5.49
episode: 1777	score: 8.0	memory length: 404990	epsilon: 0.3961178200
1019724	steps: 456	lr: 6.400000000000001e-06	evaluation reward: 5.56
episode: 1778	score: 1.0	memory length: 405162	epsilon: 0.3957772600
101951	steps: 172	lr: 6.400000000000001e-06	evaluation reward: 5.46

episode: 1779	score: 16.0	memory length: 405666	epsilon: 0.394779340
0101888	steps: 504	lr: 6.400000000000001e-06	evaluation reward:
5.58			
episode: 1780	score: 5.0	memory length: 405993	epsilon: 0.3941318800
101847	steps: 327	lr: 6.400000000000001e-06	evaluation reward: 5
.58			
episode: 1781	score: 4.0	memory length: 406290	epsilon: 0.3935438200
1018096	steps: 297	lr: 6.400000000000001e-06	evaluation reward:
5.56			
episode: 1782	score: 9.0	memory length: 406793	epsilon: 0.3925478800
1017465	steps: 503	lr: 6.400000000000001e-06	evaluation reward:
5.55			
episode: 1783	score: 2.0	memory length: 406974	epsilon: 0.3921895000
101724	steps: 181	lr: 6.400000000000001e-06	evaluation reward: 5
.54			
episode: 1784	score: 9.0	memory length: 407461	epsilon: 0.3912252400
101663	steps: 487	lr: 6.400000000000001e-06	evaluation reward: 5
.6			
episode: 1785	score: 5.0	memory length: 407785	epsilon: 0.3905837200
101622	steps: 324	lr: 6.400000000000001e-06	evaluation reward: 5
.6			
episode: 1786	score: 8.0	memory length: 408258	epsilon: 0.3896471800
101563	steps: 473	lr: 6.400000000000001e-06	evaluation reward: 5
.62			
episode: 1787	score: 4.0	memory length: 408554	epsilon: 0.3890611000
101526	steps: 296	lr: 6.400000000000001e-06	evaluation reward: 5
.58			
episode: 1788	score: 4.0	memory length: 408816	epsilon: 0.3885423400
101493	steps: 262	lr: 6.400000000000001e-06	evaluation reward: 5
.54			
episode: 1789	score: 4.0	memory length: 409093	epsilon: 0.3879938800
1014584	steps: 277	lr: 6.400000000000001e-06	evaluation reward:
5.55			
episode: 1790	score: 3.0	memory length: 409305	epsilon: 0.3875741200
101432	steps: 212	lr: 6.400000000000001e-06	evaluation reward: 5
.51			
episode: 1791	score: 5.0	memory length: 409614	epsilon: 0.3869623000
101393	steps: 309	lr: 6.400000000000001e-06	evaluation reward: 5
.52			
episode: 1792	score: 4.0	memory length: 409890	epsilon: 0.3864158200
1013586	steps: 276	lr: 6.400000000000001e-06	evaluation reward:
5.51			
episode: 1793	score: 6.0	memory length: 410265	epsilon: 0.3856733200
1013116	steps: 375	lr: 6.400000000000001e-06	evaluation reward:
5.54			
episode: 1794	score: 4.0	memory length: 410523	epsilon: 0.3851624800
1012793	steps: 258	lr: 6.400000000000001e-06	evaluation reward:
5.56			
episode: 1795	score: 6.0	memory length: 410877	epsilon: 0.3844615600
101235	steps: 354	lr: 6.400000000000001e-06	evaluation reward: 5
.58			
episode: 1796	score: 7.0	memory length: 411298	epsilon: 0.3836279800
101182	steps: 421	lr: 6.400000000000001e-06	evaluation reward: 5


```
.56
episode: 1797  score: 6.0  memory length: 411675  epsilon: 0.3828815200
101135  steps: 377  lr: 6.400000000000001e-06  evaluation reward: 5
.59
episode: 1798  score: 3.0  memory length: 411888  epsilon: 0.3824597800
1011083  steps: 213  lr: 6.400000000000001e-06  evaluation reward:
5.59
episode: 1799  score: 3.0  memory length: 412117  epsilon: 0.3820063600
1010796  steps: 229  lr: 6.400000000000001e-06  evaluation reward:
5.55
episode: 1800  score: 6.0  memory length: 412461  epsilon: 0.3813252400
1010365  steps: 344  lr: 6.400000000000001e-06  evaluation reward:
5.54
episode: 1801  score: 3.0  memory length: 412711  epsilon: 0.3808302400
101005  steps: 250  lr: 6.400000000000001e-06  evaluation reward: 5
.53
episode: 1802  score: 10.0  memory length: 413137  epsilon: 0.3799867600
0100952  steps: 426  lr: 6.400000000000001e-06  evaluation reward:
5.61
episode: 1803  score: 4.0  memory length: 413432  epsilon: 0.3794026600
100915  steps: 295  lr: 6.400000000000001e-06  evaluation reward: 5
.6
episode: 1804  score: 3.0  memory length: 413643  epsilon: 0.3789848800
1008884  steps: 211  lr: 6.400000000000001e-06  evaluation reward:
5.59
episode: 1805  score: 4.0  memory length: 413886  epsilon: 0.3785037400
100858  steps: 243  lr: 6.400000000000001e-06  evaluation reward: 5
.55
episode: 1806  score: 3.0  memory length: 414118  epsilon: 0.3780443800
100829  steps: 232  lr: 6.400000000000001e-06  evaluation reward: 5
.51
episode: 1807  score: 6.0  memory length: 414492  epsilon: 0.3773038600
100782  steps: 374  lr: 6.400000000000001e-06  evaluation reward: 5
.51
episode: 1808  score: 4.0  memory length: 414788  epsilon: 0.3767177800
100745  steps: 296  lr: 6.400000000000001e-06  evaluation reward: 5
.52
episode: 1809  score: 3.0  memory length: 415001  epsilon: 0.3762960400
1007183  steps: 213  lr: 6.400000000000001e-06  evaluation reward:
5.51
episode: 1810  score: 6.0  memory length: 415391  epsilon: 0.3755238400
1006694  steps: 390  lr: 6.400000000000001e-06  evaluation reward:
5.48
episode: 1811  score: 7.0  memory length: 415781  epsilon: 0.3747516400
1006206  steps: 390  lr: 6.400000000000001e-06  evaluation reward:
5.53
episode: 1812  score: 8.0  memory length: 416237  epsilon: 0.3738487600
1005635  steps: 456  lr: 6.400000000000001e-06  evaluation reward:
5.56
episode: 1813  score: 12.0  memory length: 416844  epsilon: 0.372646900
01004874  steps: 607  lr: 6.400000000000001e-06  evaluation reward:
5.64
episode: 1814  score: 5.0  memory length: 417152  epsilon: 0.3720370600
```

100449	steps: 308	lr: 6.400000000000001e-06	evaluation reward: 5
.68			
episode: 1815	score: 5.0	memory length: 417455	epsilon: 0.3714371200
100411	steps: 303	lr: 6.400000000000001e-06	evaluation reward: 5
.64			
episode: 1816	score: 6.0	memory length: 417808	epsilon: 0.3707381800
1003667	steps: 353	lr: 6.400000000000001e-06	evaluation reward:
5.63			
episode: 1817	score: 6.0	memory length: 418165	epsilon: 0.3700313200
100322	steps: 357	lr: 6.400000000000001e-06	evaluation reward: 5
.66			
episode: 1818	score: 4.0	memory length: 418461	epsilon: 0.3694452400
100285	steps: 296	lr: 6.400000000000001e-06	evaluation reward: 5
.66			
episode: 1819	score: 4.0	memory length: 418724	epsilon: 0.3689245000
100252	steps: 263	lr: 6.400000000000001e-06	evaluation reward: 5
.6			
episode: 1820	score: 8.0	memory length: 419164	epsilon: 0.3680533000
100197	steps: 440	lr: 6.400000000000001e-06	evaluation reward: 5
.64			
episode: 1821	score: 7.0	memory length: 419540	epsilon: 0.3673088200
1001497	steps: 376	lr: 6.400000000000001e-06	evaluation reward:
5.7			
episode: 1822	score: 8.0	memory length: 419940	epsilon: 0.3665168200
1000996	steps: 400	lr: 6.400000000000001e-06	evaluation reward:
5.71			
episode: 1823	score: 10.0	memory length: 420476	epsilon: 0.365455540
01000324	steps: 536	lr: 6.400000000000001e-06	evaluation reward:
5.73			
episode: 1824	score: 5.0	memory length: 420787	epsilon: 0.3648397600
0999935	steps: 311	lr: 6.400000000000001e-06	evaluation reward:
5.71			
episode: 1825	score: 3.0	memory length: 421016	epsilon: 0.3643863400
099965	steps: 229	lr: 6.400000000000001e-06	evaluation reward: 5
.71			
episode: 1826	score: 5.0	memory length: 421325	epsilon: 0.3637745200
099926	steps: 309	lr: 6.400000000000001e-06	evaluation reward: 5
.72			
episode: 1827	score: 3.0	memory length: 421537	epsilon: 0.3633547600
0998995	steps: 212	lr: 6.400000000000001e-06	evaluation reward:
5.69			
episode: 1828	score: 7.0	memory length: 421953	epsilon: 0.3625310800
0998474	steps: 416	lr: 6.400000000000001e-06	evaluation reward:
5.73			
episode: 1829	score: 9.0	memory length: 422344	epsilon: 0.3617569000
0997984	steps: 391	lr: 6.400000000000001e-06	evaluation reward:
5.74			
episode: 1830	score: 6.0	memory length: 422703	epsilon: 0.3610460800
0997535	steps: 359	lr: 6.400000000000001e-06	evaluation reward:
5.73			
episode: 1831	score: 7.0	memory length: 423108	epsilon: 0.3602441800
0997027	steps: 405	lr: 6.400000000000001e-06	evaluation reward:
5.73			

episode: 1832	score: 6.0	memory length: 423446	epsilon: 0.3595749400
0996604	steps: 338	lr: 6.400000000000001e-06	evaluation reward:
5.7			
episode: 1833	score: 8.0	memory length: 423910	epsilon: 0.3586562200
099602	steps: 464	lr: 6.400000000000001e-06	evaluation reward: 5
.74			
episode: 1834	score: 7.0	memory length: 424348	epsilon: 0.3577889800
0995474	steps: 438	lr: 6.400000000000001e-06	evaluation reward:
5.74			
episode: 1835	score: 8.0	memory length: 424804	epsilon: 0.3568861000
09949	steps: 456	lr: 6.400000000000001e-06	evaluation reward: 5.
79			
episode: 1836	score: 9.0	memory length: 425299	epsilon: 0.3559060000
099428	steps: 495	lr: 6.400000000000001e-06	evaluation reward: 5
.84			
episode: 1837	score: 3.0	memory length: 425545	epsilon: 0.3554189200
0993974	steps: 246	lr: 6.400000000000001e-06	evaluation reward:
5.83			
episode: 1838	score: 6.0	memory length: 425898	epsilon: 0.3547199800
099353	steps: 353	lr: 6.400000000000001e-06	evaluation reward: 5
.85			
episode: 1839	score: 7.0	memory length: 426320	epsilon: 0.3538844200
0993003	steps: 422	lr: 6.400000000000001e-06	evaluation reward:
5.87			
episode: 1840	score: 7.0	memory length: 426763	epsilon: 0.3530072800
099245	steps: 443	lr: 6.400000000000001e-06	evaluation reward: 5
.91			
episode: 1841	score: 8.0	memory length: 427182	epsilon: 0.3521776600
0991924	steps: 419	lr: 6.400000000000001e-06	evaluation reward:
5.89			
episode: 1842	score: 2.0	memory length: 427380	epsilon: 0.3517856200
0991676	steps: 198	lr: 6.400000000000001e-06	evaluation reward:
5.84			
episode: 1843	score: 8.0	memory length: 427773	epsilon: 0.3510074800
0991183	steps: 393	lr: 6.400000000000001e-06	evaluation reward:
5.85			
episode: 1844	score: 5.0	memory length: 428080	epsilon: 0.3503996200
09908	steps: 307	lr: 6.400000000000001e-06	evaluation reward: 5.
87			
episode: 1845	score: 6.0	memory length: 428480	epsilon: 0.3496076200
09903	steps: 400	lr: 6.400000000000001e-06	evaluation reward: 5.
86			
episode: 1846	score: 4.0	memory length: 428741	epsilon: 0.3490908400
098997	steps: 261	lr: 6.400000000000001e-06	evaluation reward: 5
.85			
episode: 1847	score: 7.0	memory length: 429131	epsilon: 0.3483186400
098948	steps: 390	lr: 6.400000000000001e-06	evaluation reward: 5
.89			
episode: 1848	score: 6.0	memory length: 429506	epsilon: 0.3475761400
098901	steps: 375	lr: 6.400000000000001e-06	evaluation reward: 5
.82			
episode: 1849	score: 4.0	memory length: 429769	epsilon: 0.3470554000
098868	steps: 263	lr: 6.400000000000001e-06	evaluation reward: 5

```
.78
episode: 1850  score: 2.0  memory length: 429951  epsilon: 0.3466950400
0988455      steps: 182   lr: 6.400000000000001e-06  evaluation reward:
5.77
episode: 1851  score: 6.0  memory length: 430309  epsilon: 0.3459862000
0988006      steps: 358   lr: 6.400000000000001e-06  evaluation reward:
5.7
episode: 1852  score: 4.0  memory length: 430603  epsilon: 0.3454040800
098764      steps: 294   lr: 6.400000000000001e-06  evaluation reward: 5
.68
episode: 1853  score: 4.0  memory length: 430865  epsilon: 0.3448853200
098731      steps: 262   lr: 6.400000000000001e-06  evaluation reward: 5
.68
episode: 1854  score: 5.0  memory length: 431190  epsilon: 0.3442418200
09869      steps: 325   lr: 6.400000000000001e-06  evaluation reward: 5.
68
episode: 1855  score: 5.0  memory length: 431474  epsilon: 0.3436795000
0986547     steps: 284   lr: 6.400000000000001e-06  evaluation reward:
5.69
episode: 1856  score: 6.0  memory length: 431834  epsilon: 0.3429667000
0986096     steps: 360   lr: 6.400000000000001e-06  evaluation reward:
5.68
episode: 1857  score: 8.0  memory length: 432266  epsilon: 0.3421113400
0985555     steps: 432   lr: 6.400000000000001e-06  evaluation reward:
5.69
episode: 1858  score: 2.0  memory length: 432466  epsilon: 0.3417153400
0985304     steps: 200   lr: 6.400000000000001e-06  evaluation reward:
5.68
episode: 1859  score: 6.0  memory length: 432860  epsilon: 0.3409352200
098481      steps: 394   lr: 6.400000000000001e-06  evaluation reward: 5
.7
episode: 1860  score: 4.0  memory length: 433116  epsilon: 0.3404283400
098449      steps: 256   lr: 6.400000000000001e-06  evaluation reward: 5
.67
episode: 1861  score: 6.0  memory length: 433434  epsilon: 0.3397987000
098409      steps: 318   lr: 6.400000000000001e-06  evaluation reward: 5
.69
episode: 1862  score: 4.0  memory length: 433711  epsilon: 0.3392502400
0983744     steps: 277   lr: 6.400000000000001e-06  evaluation reward:
5.68
episode: 1863  score: 8.0  memory length: 434152  epsilon: 0.3383770600
098319      steps: 441   lr: 6.400000000000001e-06  evaluation reward: 5
.7
episode: 1864  score: 6.0  memory length: 434487  epsilon: 0.3377137600
098277      steps: 335   lr: 6.400000000000001e-06  evaluation reward: 5
.7
episode: 1865  score: 3.0  memory length: 434718  epsilon: 0.3372563800
0982483     steps: 231   lr: 6.400000000000001e-06  evaluation reward:
5.67
episode: 1866  score: 3.0  memory length: 434931  epsilon: 0.3368346400
0982216     steps: 213   lr: 6.400000000000001e-06  evaluation reward:
5.63
episode: 1867  score: 11.0 memory length: 435493  epsilon: 0.335721880
```

0098151	steps: 562	lr: 6.400000000000001e-06	evaluation reward:
5.69			
episode: 1868	score: 7.0	memory length: 435868	epsilon: 0.3349793800
098104	steps: 375	lr: 6.400000000000001e-06	evaluation reward: 5
.72			
episode: 1869	score: 6.0	memory length: 436200	epsilon: 0.3343220200
0980626	steps: 332	lr: 6.400000000000001e-06	evaluation reward:
5.74			
episode: 1870	score: 4.0	memory length: 436442	epsilon: 0.3338428600
0980323	steps: 242	lr: 6.400000000000001e-06	evaluation reward:
5.72			
episode: 1871	score: 4.0	memory length: 436700	epsilon: 0.3333320200
098	steps: 258	lr: 6.400000000000001e-06	evaluation reward: 5.7
episode: 1872	score: 7.0	memory length: 437102	epsilon: 0.3325360600
0979496	steps: 402	lr: 6.400000000000001e-06	evaluation reward:
5.67			
episode: 1873	score: 8.0	memory length: 437582	epsilon: 0.3315856600
0978895	steps: 480	lr: 6.400000000000001e-06	evaluation reward:
5.71			
episode: 1874	score: 8.0	memory length: 438038	epsilon: 0.3306827800
0978324	steps: 456	lr: 6.400000000000001e-06	evaluation reward:
5.76			
episode: 1875	score: 8.0	memory length: 438425	epsilon: 0.3299165200
097784	steps: 387	lr: 6.400000000000001e-06	evaluation reward: 5
.78			
episode: 1876	score: 6.0	memory length: 438785	epsilon: 0.3292037200
097739	steps: 360	lr: 6.400000000000001e-06	evaluation reward: 5
.75			
episode: 1877	score: 6.0	memory length: 439122	epsilon: 0.3285364600
0976966	steps: 337	lr: 6.400000000000001e-06	evaluation reward:
5.73			
episode: 1878	score: 8.0	memory length: 439546	epsilon: 0.3276969400
0976435	steps: 424	lr: 6.400000000000001e-06	evaluation reward:
5.8			
episode: 1879	score: 4.0	memory length: 439806	epsilon: 0.3271821400
097611	steps: 260	lr: 6.400000000000001e-06	evaluation reward: 5
.68			
episode: 1880	score: 4.0	memory length: 440074	epsilon: 0.3266515000
0975773	steps: 268	lr: 6.400000000000001e-06	evaluation reward:
5.67			
episode: 1881	score: 4.0	memory length: 440317	epsilon: 0.3261703600
097547	steps: 243	lr: 6.400000000000001e-06	evaluation reward: 5
.67			
episode: 1882	score: 4.0	memory length: 440576	epsilon: 0.3256575400
0975144	steps: 259	lr: 6.400000000000001e-06	evaluation reward:
5.62			
episode: 1883	score: 7.0	memory length: 440969	epsilon: 0.3248794000
097465	steps: 393	lr: 6.400000000000001e-06	evaluation reward: 5
.67			
episode: 1884	score: 3.0	memory length: 441221	epsilon: 0.3243804400
0974336	steps: 252	lr: 6.400000000000001e-06	evaluation reward:
5.61			
episode: 1885	score: 9.0	memory length: 441716	epsilon: 0.3234003400

0973716	steps: 495	lr: 6.400000000000001e-06	evaluation reward:
5.65			
episode: 1886	score: 7.0	memory length: 442117	epsilon: 0.3226063600
0973214	steps: 401	lr: 6.400000000000001e-06	evaluation reward:
5.64			
episode: 1887	score: 9.0	memory length: 442567	epsilon: 0.3217153600
097265	steps: 450	lr: 6.400000000000001e-06	evaluation reward: 5
.69			
episode: 1888	score: 6.0	memory length: 442939	epsilon: 0.3209788000
0972184	steps: 372	lr: 6.400000000000001e-06	evaluation reward:
5.71			
episode: 1889	score: 9.0	memory length: 443411	epsilon: 0.3200442400
0971593	steps: 472	lr: 6.400000000000001e-06	evaluation reward:
5.76			
episode: 1890	score: 6.0	memory length: 443739	epsilon: 0.3193948000
097118	steps: 328	lr: 6.400000000000001e-06	evaluation reward: 5
.79			
episode: 1891	score: 4.0	memory length: 444002	epsilon: 0.3188740600
097085	steps: 263	lr: 6.400000000000001e-06	evaluation reward: 5
.78			
episode: 1892	score: 4.0	memory length: 444279	epsilon: 0.3183256000
0970506	steps: 277	lr: 6.400000000000001e-06	evaluation reward:
5.78			
episode: 1893	score: 6.0	memory length: 444633	epsilon: 0.3176246800
097006	steps: 354	lr: 6.400000000000001e-06	evaluation reward: 5
.78			
episode: 1894	score: 12.0	memory length: 445156	epsilon: 0.316589140
00969407	steps: 523	lr: 6.400000000000001e-06	evaluation reward:
5.86			
episode: 1895	score: 8.0	memory length: 445616	epsilon: 0.3156783400
096883	steps: 460	lr: 6.400000000000001e-06	evaluation reward: 5
.88			
episode: 1896	score: 3.0	memory length: 445842	epsilon: 0.3152308600
096855	steps: 226	lr: 6.400000000000001e-06	evaluation reward: 5
.84			
episode: 1897	score: 5.0	memory length: 446130	epsilon: 0.3146606200
0968187	steps: 288	lr: 6.400000000000001e-06	evaluation reward:
5.83			
episode: 1898	score: 8.0	memory length: 446410	epsilon: 0.3141062200
0967836	steps: 280	lr: 6.400000000000001e-06	evaluation reward:
5.88			
episode: 1899	score: 6.0	memory length: 446782	epsilon: 0.3133696600
096737	steps: 372	lr: 6.400000000000001e-06	evaluation reward: 5
.91			
episode: 1900	score: 4.0	memory length: 447038	epsilon: 0.3128627800
096705	steps: 256	lr: 6.400000000000001e-06	evaluation reward: 5
.89			
episode: 1901	score: 5.0	memory length: 447383	epsilon: 0.3121796800
0966617	steps: 345	lr: 6.400000000000001e-06	evaluation reward:
5.91			
episode: 1902	score: 9.0	memory length: 447883	epsilon: 0.3111896800
096599	steps: 500	lr: 6.400000000000001e-06	evaluation reward: 5
.9			

episode: 1903	score: 12.0	memory length: 448516	epsilon: 0.309936340
009652	steps: 633	lr: 6.400000000000001e-06	evaluation reward: 5.98
episode: 1904	score: 6.0	memory length: 448874	epsilon: 0.3092275000
0096475	steps: 358	lr: 6.400000000000001e-06	evaluation reward: 6.01
episode: 1905	score: 8.0	memory length: 449307	epsilon: 0.3083701600
00964207	steps: 433	lr: 6.400000000000001e-06	evaluation reward: 6.05
episode: 1906	score: 4.0	memory length: 449570	epsilon: 0.3078494200
0096388	steps: 263	lr: 6.400000000000001e-06	evaluation reward: 6.06
episode: 1907	score: 7.0	memory length: 449977	epsilon: 0.3070435600
0096337	steps: 407	lr: 6.400000000000001e-06	evaluation reward: 6.07
episode: 1908	score: 6.0	memory length: 450354	epsilon: 0.3062971000
00962895	steps: 377	lr: 6.400000000000001e-06	evaluation reward: 6.09
episode: 1909	score: 7.0	memory length: 450744	epsilon: 0.3055249000
00962407	steps: 390	lr: 6.400000000000001e-06	evaluation reward: 6.13
episode: 1910	score: 5.0	memory length: 451069	epsilon: 0.3048814000
00962	steps: 325	lr: 6.400000000000001e-06	evaluation reward: 6.12
episode: 1911	score: 7.0	memory length: 451439	epsilon: 0.3041488000
00961536	steps: 370	lr: 6.400000000000001e-06	evaluation reward: 6.12
episode: 1912	score: 3.0	memory length: 451652	epsilon: 0.3037270600
0096127	steps: 213	lr: 6.400000000000001e-06	evaluation reward: 6.07
episode: 1913	score: 4.0	memory length: 451952	epsilon: 0.3031330600
00960893	steps: 300	lr: 6.400000000000001e-06	evaluation reward: 5.99
episode: 1914	score: 9.0	memory length: 452458	epsilon: 0.3021311800
0096026	steps: 506	lr: 6.400000000000001e-06	evaluation reward: 6.03
episode: 1915	score: 8.0	memory length: 452844	epsilon: 0.3013669000
00959776	steps: 386	lr: 6.400000000000001e-06	evaluation reward: 6.06
episode: 1916	score: 5.0	memory length: 453152	epsilon: 0.3007570600
0095939	steps: 308	lr: 6.400000000000001e-06	evaluation reward: 6.05
episode: 1917	score: 6.0	memory length: 453526	epsilon: 0.3000165400
0095892	steps: 374	lr: 6.400000000000001e-06	evaluation reward: 6.05
episode: 1918	score: 10.0	memory length: 453996	epsilon: 0.299085940
00958333	steps: 470	lr: 6.400000000000001e-06	evaluation reward: 6.11
episode: 1919	score: 4.0	memory length: 454297	epsilon: 0.2984899600
00957956	steps: 301	lr: 6.400000000000001e-06	evaluation reward: 6.11
episode: 1920	score: 10.0	memory length: 454775	epsilon: 0.297543520
00957357	steps: 478	lr: 6.400000000000001e-06	evaluation reward:

6.13
episode: 1921 score: 10.0 memory length: 455266 epsilon: 0.296571340
0095674 steps: 491 lr: 6.400000000000001e-06 evaluation reward:
6.16
episode: 1922 score: 9.0 memory length: 455770 epsilon: 0.2955734200
0095611 steps: 504 lr: 6.400000000000001e-06 evaluation reward: 6
.17
episode: 1923 score: 9.0 memory length: 456112 epsilon: 0.2948962600
0095568 steps: 342 lr: 6.400000000000001e-06 evaluation reward: 6
.16
episode: 1924 score: 8.0 memory length: 456513 epsilon: 0.2941022800
0095518 steps: 401 lr: 6.400000000000001e-06 evaluation reward: 6
.19
episode: 1925 score: 10.0 memory length: 457099 epsilon: 0.292942000
00954446 steps: 586 lr: 6.400000000000001e-06 evaluation reward:
6.26
episode: 1926 score: 7.0 memory length: 457498 epsilon: 0.2921519800
00953946 steps: 399 lr: 6.400000000000001e-06 evaluation reward:
6.28
episode: 1927 score: 9.0 memory length: 457952 epsilon: 0.2912530600
00953377 steps: 454 lr: 6.400000000000001e-06 evaluation reward:
6.34
episode: 1928 score: 11.0 memory length: 458526 epsilon: 0.290116540
0095266 steps: 574 lr: 6.400000000000001e-06 evaluation reward:
6.38
episode: 1929 score: 5.0 memory length: 458837 epsilon: 0.2895007600
0095227 steps: 311 lr: 6.400000000000001e-06 evaluation reward: 6
.34
episode: 1930 score: 8.0 memory length: 459271 epsilon: 0.2886414400
00951725 steps: 434 lr: 6.400000000000001e-06 evaluation reward:
6.36
episode: 1931 score: 4.0 memory length: 459532 epsilon: 0.2881246600
009514 steps: 261 lr: 6.400000000000001e-06 evaluation reward: 6.
33
episode: 1932 score: 10.0 memory length: 460062 epsilon: 0.287075260
00950734 steps: 530 lr: 6.400000000000001e-06 evaluation reward:
6.37
episode: 1933 score: 7.0 memory length: 460436 epsilon: 0.2863347400
00950265 steps: 374 lr: 6.400000000000001e-06 evaluation reward:
6.36
episode: 1934 score: 11.0 memory length: 460952 epsilon: 0.285313060
0094962 steps: 516 lr: 6.400000000000001e-06 evaluation reward:
6.4
episode: 1935 score: 8.0 memory length: 461354 epsilon: 0.2845171000
00949115 steps: 402 lr: 6.400000000000001e-06 evaluation reward:
6.4
episode: 1936 score: 10.0 memory length: 461765 epsilon: 0.283703320
009486 steps: 411 lr: 6.400000000000001e-06 evaluation reward: 6
.41
episode: 1937 score: 5.0 memory length: 462074 epsilon: 0.2830915000
00948213 steps: 309 lr: 6.400000000000001e-06 evaluation reward:
6.43
episode: 1938 score: 9.0 memory length: 462531 epsilon: 0.2821866400

094764	steps: 457	lr: 6.400000000000001e-06	evaluation reward: 6
.46			
episode: 1939	score: 5.0	memory length: 462860	epsilon: 0.2815352200
094723	steps: 329	lr: 6.400000000000001e-06	evaluation reward: 6
.44			
episode: 1940	score: 5.0	memory length: 463167	epsilon: 0.2809273600
0946844	steps: 307	lr: 6.400000000000001e-06	evaluation reward:
6.42			
episode: 1941	score: 4.0	memory length: 463408	epsilon: 0.2804501800
094654	steps: 241	lr: 6.400000000000001e-06	evaluation reward: 6
.38			
episode: 1942	score: 6.0	memory length: 463783	epsilon: 0.2797076800
094607	steps: 375	lr: 6.400000000000001e-06	evaluation reward: 6
.42			
episode: 1943	score: 6.0	memory length: 464153	epsilon: 0.2789750800
094561	steps: 370	lr: 6.400000000000001e-06	evaluation reward: 6
.4			
episode: 1944	score: 4.0	memory length: 464416	epsilon: 0.2784543400
094528	steps: 263	lr: 6.400000000000001e-06	evaluation reward: 6
.39			
episode: 1945	score: 10.0	memory length: 464981	epsilon: 0.277335640
0094457	steps: 565	lr: 6.400000000000001e-06	evaluation reward:
6.43			
episode: 1946	score: 10.0	memory length: 465354	epsilon: 0.276597100
00944104	steps: 373	lr: 6.400000000000001e-06	evaluation reward:
6.49			
episode: 1947	score: 5.0	memory length: 465646	epsilon: 0.2760189400
094374	steps: 292	lr: 6.400000000000001e-06	evaluation reward: 6
.47			
episode: 1948	score: 6.0	memory length: 466021	epsilon: 0.2752764400
094327	steps: 375	lr: 6.400000000000001e-06	evaluation reward: 6
.47			
episode: 1949	score: 4.0	memory length: 466313	epsilon: 0.2746982800
0942903	steps: 292	lr: 6.400000000000001e-06	evaluation reward:
6.47			
episode: 1950	score: 7.0	memory length: 466734	epsilon: 0.2738647000
0942376	steps: 421	lr: 6.400000000000001e-06	evaluation reward:
6.52			
episode: 1951	score: 6.0	memory length: 467088	epsilon: 0.2731637800
094193	steps: 354	lr: 6.400000000000001e-06	evaluation reward: 6
.52			
episode: 1952	score: 5.0	memory length: 467434	epsilon: 0.2724787000
09415	steps: 346	lr: 6.400000000000001e-06	evaluation reward: 6.
53			
episode: 1953	score: 6.0	memory length: 467812	epsilon: 0.2717302600
0941025	steps: 378	lr: 6.400000000000001e-06	evaluation reward:
6.55			
episode: 1954	score: 7.0	memory length: 468170	epsilon: 0.2710214200
0940577	steps: 358	lr: 6.400000000000001e-06	evaluation reward:
6.57			
episode: 1955	score: 11.0	memory length: 468743	epsilon: 0.269886880
0093986	steps: 573	lr: 6.400000000000001e-06	evaluation reward:
6.63			

episode: 1956	score: 4.0	memory length: 469021	epsilon: 0.2693364400
093951	steps: 278	lr: 6.400000000000001e-06	evaluation reward: 6
.61			
episode: 1957	score: 7.0	memory length: 469411	epsilon: 0.2685642400
093902	steps: 390	lr: 6.400000000000001e-06	evaluation reward: 6
.6			
episode: 1958	score: 5.0	memory length: 469737	epsilon: 0.2679187600
0938614	steps: 326	lr: 6.400000000000001e-06	evaluation reward:
6.63			
episode: 1959	score: 7.0	memory length: 470106	epsilon: 0.2671881400
093815	steps: 369	lr: 6.400000000000001e-06	evaluation reward: 6
.64			
episode: 1960	score: 6.0	memory length: 470409	epsilon: 0.2665882000
093777	steps: 303	lr: 6.400000000000001e-06	evaluation reward: 6
.66			
episode: 1961	score: 6.0	memory length: 470768	epsilon: 0.2658773800
093732	steps: 359	lr: 6.400000000000001e-06	evaluation reward: 6
.66			
episode: 1962	score: 4.0	memory length: 471027	epsilon: 0.2653645600
0937	steps: 259	lr: 6.400000000000001e-06	evaluation reward: 6.6
6			
episode: 1963	score: 8.0	memory length: 471416	epsilon: 0.2645943400
093651	steps: 389	lr: 6.400000000000001e-06	evaluation reward: 6
.66			
episode: 1964	score: 7.0	memory length: 471804	epsilon: 0.2638261000
0936024	steps: 388	lr: 6.400000000000001e-06	evaluation reward:
6.67			
episode: 1965	score: 5.0	memory length: 472111	epsilon: 0.2632182400
093564	steps: 307	lr: 6.400000000000001e-06	evaluation reward: 6
.69			
episode: 1966	score: 11.0	memory length: 472653	epsilon: 0.262145080
0093496	steps: 542	lr: 6.400000000000001e-06	evaluation reward:
6.77			
episode: 1967	score: 10.0	memory length: 473202	epsilon: 0.261058060
00934273	steps: 549	lr: 6.400000000000001e-06	evaluation reward:
6.76			
episode: 1968	score: 4.0	memory length: 473464	epsilon: 0.2605393000
0933945	steps: 262	lr: 6.400000000000001e-06	evaluation reward:
6.73			
episode: 1969	score: 9.0	memory length: 473975	epsilon: 0.2595275200
0933305	steps: 511	lr: 6.400000000000001e-06	evaluation reward:
6.76			
episode: 1970	score: 5.0	memory length: 474284	epsilon: 0.2589157000
093292	steps: 309	lr: 6.400000000000001e-06	evaluation reward: 6
.77			
episode: 1971	score: 10.0	memory length: 474791	epsilon: 0.257911840
0093228	steps: 507	lr: 6.400000000000001e-06	evaluation reward:
6.83			
episode: 1972	score: 4.0	memory length: 475034	epsilon: 0.2574307000
093198	steps: 243	lr: 6.400000000000001e-06	evaluation reward: 6
.8			
episode: 1973	score: 7.0	memory length: 475458	epsilon: 0.2565911800
0931447	steps: 424	lr: 6.400000000000001e-06	evaluation reward:

```
6.79
episode: 1974  score: 12.0  memory length: 476029  epsilon: 0.255460600
0093073  steps: 571  lr: 6.400000000000001e-06  evaluation reward:
6.83
episode: 1975  score: 8.0  memory length: 476461  epsilon: 0.2546052400
093019  steps: 432  lr: 6.400000000000001e-06  evaluation reward: 6
.83
episode: 1976  score: 3.0  memory length: 476691  epsilon: 0.2541498400
09299  steps: 230  lr: 6.400000000000001e-06  evaluation reward: 6.
8
episode: 1977  score: 11.0  memory length: 477246  epsilon: 0.253050940
00929207  steps: 555  lr: 6.400000000000001e-06  evaluation reward:
6.85
episode: 1978  score: 5.0  memory length: 477551  epsilon: 0.2524470400
0928825  steps: 305  lr: 6.400000000000001e-06  evaluation reward:
6.82
episode: 1979  score: 5.0  memory length: 477855  epsilon: 0.2518451200
0928444  steps: 304  lr: 6.400000000000001e-06  evaluation reward:
6.83
episode: 1980  score: 6.0  memory length: 478246  epsilon: 0.2510709400
0927954  steps: 391  lr: 6.400000000000001e-06  evaluation reward:
6.85
episode: 1981  score: 11.0  memory length: 478832  epsilon: 0.249910660
0092722  steps: 586  lr: 6.400000000000001e-06  evaluation reward:
6.92
episode: 1982  score: 10.0  memory length: 479360  epsilon: 0.248865220
0092656  steps: 528  lr: 6.400000000000001e-06  evaluation reward:
6.98
episode: 1983  score: 9.0  memory length: 479848  epsilon: 0.2478989800
0925947  steps: 488  lr: 6.400000000000001e-06  evaluation reward:
7.0
episode: 1984  score: 5.0  memory length: 480139  epsilon: 0.2473228000
0925583  steps: 291  lr: 6.400000000000001e-06  evaluation reward:
7.02
episode: 1985  score: 11.0  memory length: 480693  epsilon: 0.246225880
0092489  steps: 554  lr: 6.400000000000001e-06  evaluation reward:
7.04
episode: 1986  score: 9.0  memory length: 481148  epsilon: 0.2453249800
092432  steps: 455  lr: 6.400000000000001e-06  evaluation reward: 7
.06
episode: 1987  score: 4.0  memory length: 481407  epsilon: 0.2448121600
0923994  steps: 259  lr: 6.400000000000001e-06  evaluation reward:
7.01
episode: 1988  score: 10.0  memory length: 481889  epsilon: 0.243857800
0092339  steps: 482  lr: 6.400000000000001e-06  evaluation reward:
7.05
episode: 1989  score: 3.0  memory length: 482136  epsilon: 0.2433687400
092308  steps: 247  lr: 6.400000000000001e-06  evaluation reward: 6
.99
episode: 1990  score: 4.0  memory length: 482433  epsilon: 0.2427806800
092271  steps: 297  lr: 6.400000000000001e-06  evaluation reward: 6
.97
episode: 1991  score: 7.0  memory length: 482857  epsilon: 0.2419411600
```

0922178	steps: 424	lr: 6.400000000000001e-06	evaluation reward:
7.0			
episode: 1992	score: 10.0	memory length: 483393	epsilon: 0.240879880
00921506	steps: 536	lr: 6.400000000000001e-06	evaluation reward:
7.06			
episode: 1993	score: 8.0	memory length: 483827	epsilon: 0.2400205600
0920963	steps: 434	lr: 6.400000000000001e-06	evaluation reward:
7.08			
episode: 1994	score: 7.0	memory length: 484253	epsilon: 0.2391770800
092043	steps: 426	lr: 6.400000000000001e-06	evaluation reward: 7
.03			
episode: 1995	score: 8.0	memory length: 484645	epsilon: 0.2384009200
0919938	steps: 392	lr: 6.400000000000001e-06	evaluation reward:
7.03			
episode: 1996	score: 6.0	memory length: 485017	epsilon: 0.2376643600
0919472	steps: 372	lr: 6.400000000000001e-06	evaluation reward:
7.06			
episode: 1997	score: 6.0	memory length: 485354	epsilon: 0.2369971000
091905	steps: 337	lr: 6.400000000000001e-06	evaluation reward: 7
.07			
episode: 1998	score: 11.0	memory length: 485897	epsilon: 0.235921960
0091837	steps: 543	lr: 6.400000000000001e-06	evaluation reward:
7.1			
episode: 1999	score: 11.0	memory length: 486405	epsilon: 0.234916120
00917733	steps: 508	lr: 6.400000000000001e-06	evaluation reward:
7.15			
episode: 2000	score: 7.0	memory length: 486771	epsilon: 0.2341914400
0917275	steps: 366	lr: 6.400000000000001e-06	evaluation reward:
7.18			
episode: 2001	score: 4.0	memory length: 487049	epsilon: 0.2336410000
0916926	steps: 278	lr: 6.400000000000001e-06	evaluation reward:
7.17			
episode: 2002	score: 7.0	memory length: 487436	epsilon: 0.2328747400
0916442	steps: 387	lr: 6.400000000000001e-06	evaluation reward:
7.15			
episode: 2003	score: 5.0	memory length: 487746	epsilon: 0.2322609400
0916053	steps: 310	lr: 6.400000000000001e-06	evaluation reward:
7.08			
episode: 2004	score: 6.0	memory length: 488072	epsilon: 0.2316154600
0915645	steps: 326	lr: 6.400000000000001e-06	evaluation reward:
7.08			
episode: 2005	score: 11.0	memory length: 488625	epsilon: 0.230520520
00914952	steps: 553	lr: 6.400000000000001e-06	evaluation reward:
7.11			
episode: 2006	score: 4.0	memory length: 488900	epsilon: 0.2299760200
0914608	steps: 275	lr: 6.400000000000001e-06	evaluation reward:
7.11			
episode: 2007	score: 7.0	memory length: 489303	epsilon: 0.2291780800
0914103	steps: 403	lr: 6.400000000000001e-06	evaluation reward:
7.11			
episode: 2008	score: 5.0	memory length: 489584	epsilon: 0.2286217000
091375	steps: 281	lr: 6.400000000000001e-06	evaluation reward: 7
.1			

episode: 2009	score: 8.0	memory length: 489992	epsilon: 0.2278138600
091324	steps: 408	lr: 6.400000000000001e-06	evaluation reward: 7
.11			
episode: 2010	score: 5.0	memory length: 490320	epsilon: 0.2271644200
091283	steps: 328	lr: 6.400000000000001e-06	evaluation reward: 7
.11			
episode: 2011	score: 5.0	memory length: 490632	epsilon: 0.2265466600
0912438	steps: 312	lr: 6.400000000000001e-06	evaluation reward:
7.09			
episode: 2012	score: 4.0	memory length: 490876	epsilon: 0.2260635400
0912132	steps: 244	lr: 6.400000000000001e-06	evaluation reward:
7.1			
episode: 2013	score: 7.0	memory length: 491319	epsilon: 0.2251864000
0911577	steps: 443	lr: 6.400000000000001e-06	evaluation reward:
7.13			
episode: 2014	score: 8.0	memory length: 491744	epsilon: 0.2243449000
0911045	steps: 425	lr: 6.400000000000001e-06	evaluation reward:
7.12			
episode: 2015	score: 4.0	memory length: 492039	epsilon: 0.2237608000
0910675	steps: 295	lr: 6.400000000000001e-06	evaluation reward:
7.08			
episode: 2016	score: 5.0	memory length: 492367	epsilon: 0.2231113600
0910264	steps: 328	lr: 6.400000000000001e-06	evaluation reward:
7.08			
episode: 2017	score: 5.0	memory length: 492693	epsilon: 0.2224658800
0909856	steps: 326	lr: 6.400000000000001e-06	evaluation reward:
7.07			
episode: 2018	score: 8.0	memory length: 493148	epsilon: 0.2215649800
0909286	steps: 455	lr: 6.400000000000001e-06	evaluation reward:
7.05			
episode: 2019	score: 12.0	memory length: 493741	epsilon: 0.220390840
00908543	steps: 593	lr: 6.400000000000001e-06	evaluation reward:
7.13			
episode: 2020	score: 8.0	memory length: 494175	epsilon: 0.2195315200
0908	steps: 434	lr: 6.400000000000001e-06	evaluation reward: 7.1
1			
episode: 2021	score: 5.0	memory length: 494448	epsilon: 0.2189909800
0907657	steps: 273	lr: 6.400000000000001e-06	evaluation reward:
7.06			
episode: 2022	score: 9.0	memory length: 494927	epsilon: 0.2180425600
0907057	steps: 479	lr: 6.400000000000001e-06	evaluation reward:
7.06			
episode: 2023	score: 7.0	memory length: 495347	epsilon: 0.2172109600
090653	steps: 420	lr: 6.400000000000001e-06	evaluation reward: 7
.04			
episode: 2024	score: 6.0	memory length: 495736	epsilon: 0.2164407400
0906044	steps: 389	lr: 6.400000000000001e-06	evaluation reward:
7.02			
episode: 2025	score: 9.0	memory length: 496203	epsilon: 0.2155160800
090546	steps: 467	lr: 6.400000000000001e-06	evaluation reward: 7
.01			
episode: 2026	score: 9.0	memory length: 496683	epsilon: 0.2145656800
0904858	steps: 480	lr: 6.400000000000001e-06	evaluation reward:

7.03
episode: 2027 score: 11.0 memory length: 497235 epsilon: 0.213472720
00904166 steps: 552 lr: 6.400000000000001e-06 evaluation reward:
7.05
episode: 2028 score: 6.0 memory length: 497610 epsilon: 0.2127302200
0903696 steps: 375 lr: 6.400000000000001e-06 evaluation reward:
7.0
episode: 2029 score: 6.0 memory length: 497947 epsilon: 0.2120629600
0903274 steps: 337 lr: 6.400000000000001e-06 evaluation reward:
7.01
episode: 2030 score: 9.0 memory length: 498434 epsilon: 0.2110987000
0902664 steps: 487 lr: 6.400000000000001e-06 evaluation reward:
7.02
episode: 2031 score: 6.0 memory length: 498771 epsilon: 0.2104314400
0902242 steps: 337 lr: 6.400000000000001e-06 evaluation reward:
7.04
episode: 2032 score: 4.0 memory length: 499065 epsilon: 0.2098493200
0901874 steps: 294 lr: 6.400000000000001e-06 evaluation reward:
6.98
episode: 2033 score: 4.0 memory length: 499342 epsilon: 0.2093008600
0901527 steps: 277 lr: 6.400000000000001e-06 evaluation reward:
6.95
episode: 2034 score: 4.0 memory length: 499603 epsilon: 0.2087840800
09012 steps: 261 lr: 6.400000000000001e-06 evaluation reward: 6.
88
episode: 2035 score: 7.0 memory length: 500022 epsilon: 0.2079544600
0900675 steps: 419 lr: 2.560000000000001e-06 evaluation reward:
6.87
episode: 2036 score: 8.0 memory length: 500494 epsilon: 0.2070199000
0900083 steps: 472 lr: 2.560000000000001e-06 evaluation reward:
6.85
episode: 2037 score: 7.0 memory length: 500919 epsilon: 0.2061784000
089955 steps: 425 lr: 2.560000000000001e-06 evaluation reward: 6
.87
episode: 2038 score: 8.0 memory length: 501387 epsilon: 0.2052517600
0898965 steps: 468 lr: 2.560000000000001e-06 evaluation reward:
6.86
episode: 2039 score: 5.0 memory length: 501673 epsilon: 0.2046854800
0898606 steps: 286 lr: 2.560000000000001e-06 evaluation reward:
6.86
episode: 2040 score: 4.0 memory length: 501933 epsilon: 0.2041706800
089828 steps: 260 lr: 2.560000000000001e-06 evaluation reward: 6
.85
episode: 2041 score: 8.0 memory length: 502380 epsilon: 0.2032856200
089772 steps: 447 lr: 2.560000000000001e-06 evaluation reward: 6
.89
episode: 2042 score: 7.0 memory length: 502746 epsilon: 0.2025609400
0897262 steps: 366 lr: 2.560000000000001e-06 evaluation reward:
6.9
episode: 2043 score: 5.0 memory length: 503092 epsilon: 0.2018758600
089683 steps: 346 lr: 2.560000000000001e-06 evaluation reward: 6
.89
episode: 2044 score: 8.0 memory length: 503530 epsilon: 0.2010086200

089628	steps: 438	lr: 2.560000000000001e-06	evaluation reward: 6
.93			
episode: 2045	score: 4.0	memory length: 503790	epsilon: 0.2004938200
0895954	steps: 260	lr: 2.560000000000001e-06	evaluation reward:
6.87			
episode: 2046	score: 9.0	memory length: 504288	epsilon: 0.1995077800
089533	steps: 498	lr: 2.560000000000001e-06	evaluation reward: 6
.86			
episode: 2047	score: 8.0	memory length: 504748	epsilon: 0.1985969800
0894754	steps: 460	lr: 2.560000000000001e-06	evaluation reward:
6.89			
episode: 2048	score: 8.0	memory length: 505203	epsilon: 0.1976960800
0894184	steps: 455	lr: 2.560000000000001e-06	evaluation reward:
6.91			
episode: 2049	score: 3.0	memory length: 505416	epsilon: 0.1972743400
0893918	steps: 213	lr: 2.560000000000001e-06	evaluation reward:
6.9			
episode: 2050	score: 9.0	memory length: 505881	epsilon: 0.1963536400
0893335	steps: 465	lr: 2.560000000000001e-06	evaluation reward:
6.92			
episode: 2051	score: 8.0	memory length: 506315	epsilon: 0.1954943200
089279	steps: 434	lr: 2.560000000000001e-06	evaluation reward: 6
.94			
episode: 2052	score: 8.0	memory length: 506772	epsilon: 0.1945894600
089222	steps: 457	lr: 2.560000000000001e-06	evaluation reward: 6
.97			
episode: 2053	score: 7.0	memory length: 507173	epsilon: 0.1937954800
0891716	steps: 401	lr: 2.560000000000001e-06	evaluation reward:
6.98			
episode: 2054	score: 4.0	memory length: 507452	epsilon: 0.1932430600
0891367	steps: 279	lr: 2.560000000000001e-06	evaluation reward:
6.95			
episode: 2055	score: 6.0	memory length: 507790	epsilon: 0.1925738200
0890944	steps: 338	lr: 2.560000000000001e-06	evaluation reward:
6.9			
episode: 2056	score: 5.0	memory length: 508114	epsilon: 0.1919323000
0890538	steps: 324	lr: 2.560000000000001e-06	evaluation reward:
6.91			
episode: 2057	score: 6.0	memory length: 508485	epsilon: 0.1911977200
0890073	steps: 371	lr: 2.560000000000001e-06	evaluation reward:
6.9			
episode: 2058	score: 8.0	memory length: 508928	epsilon: 0.1903205800
0889518	steps: 443	lr: 2.560000000000001e-06	evaluation reward:
6.93			
episode: 2059	score: 7.0	memory length: 509332	epsilon: 0.1895206600
0889012	steps: 404	lr: 2.560000000000001e-06	evaluation reward:
6.93			
episode: 2060	score: 5.0	memory length: 509623	epsilon: 0.1889444800
0888647	steps: 291	lr: 2.560000000000001e-06	evaluation reward:
6.92			
episode: 2061	score: 9.0	memory length: 510159	epsilon: 0.1878832000
0887976	steps: 536	lr: 2.560000000000001e-06	evaluation reward:
6.95			

episode: 2062	score: 4.0	memory length: 510416	epsilon: 0.1873743400
0887654	steps: 257	lr: 2.560000000000001e-06	evaluation reward:
6.95			
episode: 2063	score: 3.0	memory length: 510629	epsilon: 0.1869526000
0887387	steps: 213	lr: 2.560000000000001e-06	evaluation reward:
6.9			
episode: 2064	score: 8.0	memory length: 511089	epsilon: 0.1860418000
088681	steps: 460	lr: 2.560000000000001e-06	evaluation reward: 6
.91			
episode: 2065	score: 7.0	memory length: 511485	epsilon: 0.1852577200
0886315	steps: 396	lr: 2.560000000000001e-06	evaluation reward:
6.93			
episode: 2066	score: 3.0	memory length: 511714	epsilon: 0.1848043000
0886028	steps: 229	lr: 2.560000000000001e-06	evaluation reward:
6.85			
episode: 2067	score: 6.0	memory length: 512033	epsilon: 0.1841726800
0885628	steps: 319	lr: 2.560000000000001e-06	evaluation reward:
6.81			
episode: 2068	score: 4.0	memory length: 512277	epsilon: 0.1836895600
0885323	steps: 244	lr: 2.560000000000001e-06	evaluation reward:
6.81			
episode: 2069	score: 9.0	memory length: 512816	epsilon: 0.1826223400
0884647	steps: 539	lr: 2.560000000000001e-06	evaluation reward:
6.81			
episode: 2070	score: 8.0	memory length: 513273	epsilon: 0.1817174800
0884075	steps: 457	lr: 2.560000000000001e-06	evaluation reward:
6.84			
episode: 2071	score: 7.0	memory length: 513622	epsilon: 0.1810264600
0883638	steps: 349	lr: 2.560000000000001e-06	evaluation reward:
6.81			
episode: 2072	score: 6.0	memory length: 513999	epsilon: 0.1802800000
0883165	steps: 377	lr: 2.560000000000001e-06	evaluation reward:
6.83			
episode: 2073	score: 6.0	memory length: 514370	epsilon: 0.1795454200
08827	steps: 371	lr: 2.560000000000001e-06	evaluation reward: 6.
82			
episode: 2074	score: 6.0	memory length: 514694	epsilon: 0.1789039000
0882295	steps: 324	lr: 2.560000000000001e-06	evaluation reward:
6.76			
episode: 2075	score: 2.0	memory length: 514876	epsilon: 0.1785435400
0882067	steps: 182	lr: 2.560000000000001e-06	evaluation reward:
6.7			
episode: 2076	score: 8.0	memory length: 515320	epsilon: 0.1776644200
088151	steps: 444	lr: 2.560000000000001e-06	evaluation reward: 6
.75			
episode: 2077	score: 8.0	memory length: 515718	epsilon: 0.1768763800
0881012	steps: 398	lr: 2.560000000000001e-06	evaluation reward:
6.72			
episode: 2078	score: 12.0	memory length: 516306	epsilon: 0.175712140
00880275	steps: 588	lr: 2.560000000000001e-06	evaluation reward:
6.79			
episode: 2079	score: 6.0	memory length: 516647	epsilon: 0.1750369600
0879848	steps: 341	lr: 2.560000000000001e-06	evaluation reward:

6.8
episode: 2080 score: 7.0 memory length: 517015 epsilon: 0.1743083200
0879387 steps: 368 lr: 2.560000000000001e-06 evaluation reward:
6.81
episode: 2081 score: 13.0 memory length: 517550 epsilon: 0.173249020
00878717 steps: 535 lr: 2.560000000000001e-06 evaluation reward:
6.83
episode: 2082 score: 8.0 memory length: 517846 epsilon: 0.1726629400
0878346 steps: 296 lr: 2.560000000000001e-06 evaluation reward:
6.81
episode: 2083 score: 10.0 memory length: 518353 epsilon: 0.171659080
0087771 steps: 507 lr: 2.560000000000001e-06 evaluation reward:
6.82
episode: 2084 score: 4.0 memory length: 518633 epsilon: 0.1711046800
087736 steps: 280 lr: 2.560000000000001e-06 evaluation reward: 6
.81
episode: 2085 score: 12.0 memory length: 519143 epsilon: 0.170094880
0087672 steps: 510 lr: 2.560000000000001e-06 evaluation reward:
6.82
episode: 2086 score: 6.0 memory length: 519490 epsilon: 0.1694078200
0876287 steps: 347 lr: 2.560000000000001e-06 evaluation reward:
6.79
episode: 2087 score: 8.0 memory length: 519914 epsilon: 0.1685683000
0875755 steps: 424 lr: 2.560000000000001e-06 evaluation reward:
6.83
episode: 2088 score: 3.0 memory length: 520160 epsilon: 0.1680812200
0875447 steps: 246 lr: 2.560000000000001e-06 evaluation reward:
6.76
episode: 2089 score: 11.0 memory length: 520690 epsilon: 0.167031820
00874783 steps: 530 lr: 2.560000000000001e-06 evaluation reward:
6.84
episode: 2090 score: 3.0 memory length: 520901 epsilon: 0.1666140400
087452 steps: 211 lr: 2.560000000000001e-06 evaluation reward: 6
.83
episode: 2091 score: 8.0 memory length: 521299 epsilon: 0.1658260000
087402 steps: 398 lr: 2.560000000000001e-06 evaluation reward: 6
.84
episode: 2092 score: 8.0 memory length: 521728 epsilon: 0.1649765800
0873483 steps: 429 lr: 2.560000000000001e-06 evaluation reward:
6.82
episode: 2093 score: 4.0 memory length: 522005 epsilon: 0.1644281200
0873136 steps: 277 lr: 2.560000000000001e-06 evaluation reward:
6.78
episode: 2094 score: 8.0 memory length: 522426 epsilon: 0.1635945400
0872609 steps: 421 lr: 2.560000000000001e-06 evaluation reward:
6.79
episode: 2095 score: 6.0 memory length: 522754 epsilon: 0.1629451000
0872198 steps: 328 lr: 2.560000000000001e-06 evaluation reward:
6.77
episode: 2096 score: 8.0 memory length: 523170 epsilon: 0.1621214200
0871677 steps: 416 lr: 2.560000000000001e-06 evaluation reward:
6.79
episode: 2097 score: 8.0 memory length: 523593 epsilon: 0.1612838800

0871147	steps: 423	lr: 2.560000000000001e-06	evaluation reward:
6.81			
episode: 2098	score: 6.0	memory length: 523949	epsilon: 0.1605790000
08707	steps: 356	lr: 2.560000000000001e-06	evaluation reward: 6.
76			
episode: 2099	score: 6.0	memory length: 524304	epsilon: 0.1598761000
0870256	steps: 355	lr: 2.560000000000001e-06	evaluation reward:
6.71			
episode: 2100	score: 4.0	memory length: 524579	epsilon: 0.1593316000
0869911	steps: 275	lr: 2.560000000000001e-06	evaluation reward:
6.68			
episode: 2101	score: 7.0	memory length: 524948	epsilon: 0.1586009800
086945	steps: 369	lr: 2.560000000000001e-06	evaluation reward: 6
.71			
episode: 2102	score: 5.0	memory length: 525294	epsilon: 0.1579159000
0869016	steps: 346	lr: 2.560000000000001e-06	evaluation reward:
6.69			
episode: 2103	score: 9.0	memory length: 525737	epsilon: 0.1570387600
086846	steps: 443	lr: 2.560000000000001e-06	evaluation reward: 6
.73			
episode: 2104	score: 4.0	memory length: 525998	epsilon: 0.1565219800
0868134	steps: 261	lr: 2.560000000000001e-06	evaluation reward:
6.71			
episode: 2105	score: 9.0	memory length: 526426	epsilon: 0.1556745400
0867598	steps: 428	lr: 2.560000000000001e-06	evaluation reward:
6.69			
episode: 2106	score: 6.0	memory length: 526744	epsilon: 0.1550449000
08672	steps: 318	lr: 2.560000000000001e-06	evaluation reward: 6.
71			
episode: 2107	score: 6.0	memory length: 527099	epsilon: 0.1543420000
0866755	steps: 355	lr: 2.560000000000001e-06	evaluation reward:
6.7			
episode: 2108	score: 13.0	memory length: 527717	epsilon: 0.153118360
0086598	steps: 618	lr: 2.560000000000001e-06	evaluation reward:
6.78			
episode: 2109	score: 8.0	memory length: 528189	epsilon: 0.1521838000
086539	steps: 472	lr: 2.560000000000001e-06	evaluation reward: 6
.78			
episode: 2110	score: 7.0	memory length: 528609	epsilon: 0.1513522000
0864863	steps: 420	lr: 2.560000000000001e-06	evaluation reward:
6.8			
episode: 2111	score: 7.0	memory length: 529011	epsilon: 0.1505562400
086436	steps: 402	lr: 2.560000000000001e-06	evaluation reward: 6
.82			
episode: 2112	score: 8.0	memory length: 529436	epsilon: 0.1497147400
0863827	steps: 425	lr: 2.560000000000001e-06	evaluation reward:
6.86			
episode: 2113	score: 5.0	memory length: 529744	epsilon: 0.1491049000
086344	steps: 308	lr: 2.560000000000001e-06	evaluation reward: 6
.84			
episode: 2114	score: 9.0	memory length: 530268	epsilon: 0.1480673800
0862785	steps: 524	lr: 2.560000000000001e-06	evaluation reward:
6.85			

episode: 2115	score: 5.0	memory length: 530559	epsilon: 0.1474912000
086242	steps: 291	lr: 2.560000000000001e-06	evaluation reward: 6.86
episode: 2116	score: 10.0	memory length: 530953	epsilon: 0.146711080
00861927	steps: 394	lr: 2.560000000000001e-06	evaluation reward: 6.91
episode: 2117	score: 4.0	memory length: 531250	epsilon: 0.1461230200
0861555	steps: 297	lr: 2.560000000000001e-06	evaluation reward: 6.9
episode: 2118	score: 6.0	memory length: 531609	epsilon: 0.1454122000
0861105	steps: 359	lr: 2.560000000000001e-06	evaluation reward: 6.88
episode: 2119	score: 5.0	memory length: 531939	epsilon: 0.1447588000
086069	steps: 330	lr: 2.560000000000001e-06	evaluation reward: 6.81
episode: 2120	score: 5.0	memory length: 532261	epsilon: 0.1441212400
0860288	steps: 322	lr: 2.560000000000001e-06	evaluation reward: 6.78
episode: 2121	score: 5.0	memory length: 532551	epsilon: 0.1435470400
0859925	steps: 290	lr: 2.560000000000001e-06	evaluation reward: 6.78
episode: 2122	score: 7.0	memory length: 532939	epsilon: 0.1427788000
085944	steps: 388	lr: 2.560000000000001e-06	evaluation reward: 6.76
episode: 2123	score: 4.0	memory length: 533197	epsilon: 0.1422679600
0859115	steps: 258	lr: 2.560000000000001e-06	evaluation reward: 6.73
episode: 2124	score: 7.0	memory length: 533599	epsilon: 0.1414720000
0858612	steps: 402	lr: 2.560000000000001e-06	evaluation reward: 6.74
episode: 2125	score: 4.0	memory length: 533862	epsilon: 0.1409512600
0858282	steps: 263	lr: 2.560000000000001e-06	evaluation reward: 6.69
episode: 2126	score: 8.0	memory length: 534293	epsilon: 0.1400978800
0857742	steps: 431	lr: 2.560000000000001e-06	evaluation reward: 6.68
episode: 2127	score: 10.0	memory length: 534808	epsilon: 0.139078180
00857097	steps: 515	lr: 2.560000000000001e-06	evaluation reward: 6.67
episode: 2128	score: 13.0	memory length: 535297	epsilon: 0.138109960
00856485	steps: 489	lr: 2.560000000000001e-06	evaluation reward: 6.74
episode: 2129	score: 4.0	memory length: 535556	epsilon: 0.1375971400
085616	steps: 259	lr: 2.560000000000001e-06	evaluation reward: 6.72
episode: 2130	score: 12.0	memory length: 536018	epsilon: 0.136682380
00855581	steps: 462	lr: 2.560000000000001e-06	evaluation reward: 6.75
episode: 2131	score: 6.0	memory length: 536373	epsilon: 0.1359794800
0855137	steps: 355	lr: 2.560000000000001e-06	evaluation reward: 6.75
episode: 2132	score: 9.0	memory length: 536857	epsilon: 0.1350211600
085453	steps: 484	lr: 2.560000000000001e-06	evaluation reward: 6

.8
episode: 2133 score: 6.0 memory length: 537196 epsilon: 0.1343499400
0854106 steps: 339 lr: 2.560000000000001e-06 evaluation reward:
6.82
episode: 2134 score: 13.0 memory length: 537810 epsilon: 0.133134220
00853337 steps: 614 lr: 2.560000000000001e-06 evaluation reward:
6.91
episode: 2135 score: 6.0 memory length: 538185 epsilon: 0.1323917200
0852867 steps: 375 lr: 2.560000000000001e-06 evaluation reward:
6.9
episode: 2136 score: 7.0 memory length: 538592 epsilon: 0.1315858600
0852357 steps: 407 lr: 2.560000000000001e-06 evaluation reward:
6.89
episode: 2137 score: 8.0 memory length: 539037 epsilon: 0.1307047600
08518 steps: 445 lr: 2.560000000000001e-06 evaluation reward: 6.
9
episode: 2138 score: 8.0 memory length: 539492 epsilon: 0.1298038600
085123 steps: 455 lr: 2.560000000000001e-06 evaluation reward: 6
.9
episode: 2139 score: 8.0 memory length: 539971 epsilon: 0.1288554400
085063 steps: 479 lr: 2.560000000000001e-06 evaluation reward: 6
.93
episode: 2140 score: 9.0 memory length: 540436 epsilon: 0.1279347400
0850047 steps: 465 lr: 2.560000000000001e-06 evaluation reward:
6.98
episode: 2141 score: 9.0 memory length: 540918 epsilon: 0.1269803800
0849443 steps: 482 lr: 2.560000000000001e-06 evaluation reward:
6.99
episode: 2142 score: 11.0 memory length: 541391 epsilon: 0.126043840
0084885 steps: 473 lr: 2.560000000000001e-06 evaluation reward:
7.03
episode: 2143 score: 12.0 memory length: 541824 epsilon: 0.125186500
00848308 steps: 433 lr: 2.560000000000001e-06 evaluation reward:
7.1
episode: 2144 score: 15.0 memory length: 542451 epsilon: 0.123945040
00848262 steps: 627 lr: 2.560000000000001e-06 evaluation reward:
7.17
episode: 2145 score: 5.0 memory length: 542774 epsilon: 0.1233055000
0848306 steps: 323 lr: 2.560000000000001e-06 evaluation reward:
7.18
episode: 2146 score: 6.0 memory length: 543130 epsilon: 0.1226006200
0848354 steps: 356 lr: 2.560000000000001e-06 evaluation reward:
7.15
episode: 2147 score: 5.0 memory length: 543441 epsilon: 0.1219848400
0848396 steps: 311 lr: 2.560000000000001e-06 evaluation reward:
7.12
episode: 2148 score: 10.0 memory length: 543862 epsilon: 0.121151260
00848453 steps: 421 lr: 2.560000000000001e-06 evaluation reward:
7.14
episode: 2149 score: 7.0 memory length: 544229 epsilon: 0.1204246000
0848502 steps: 367 lr: 2.560000000000001e-06 evaluation reward:
7.18
episode: 2150 score: 7.0 memory length: 544619 epsilon: 0.1196524000

0848555	steps: 390	lr: 2.560000000000001e-06	evaluation reward:
7.16			
episode: 2151	score: 6.0	memory length: 544993	epsilon: 0.1189118800
0848606	steps: 374	lr: 2.560000000000001e-06	evaluation reward:
7.14			
episode: 2152	score: 4.0	memory length: 545271	epsilon: 0.1183614400
0848643	steps: 278	lr: 2.560000000000001e-06	evaluation reward:
7.1			
episode: 2153	score: 8.0	memory length: 545747	epsilon: 0.1174189600
0848707	steps: 476	lr: 2.560000000000001e-06	evaluation reward:
7.11			
episode: 2154	score: 4.0	memory length: 546044	epsilon: 0.1168309000
0848748	steps: 297	lr: 2.560000000000001e-06	evaluation reward:
7.11			
episode: 2155	score: 9.0	memory length: 546532	epsilon: 0.1158646600
0848814	steps: 488	lr: 2.560000000000001e-06	evaluation reward:
7.14			
episode: 2156	score: 9.0	memory length: 547013	epsilon: 0.1149122800
0848878	steps: 481	lr: 2.560000000000001e-06	evaluation reward:
7.18			
episode: 2157	score: 7.0	memory length: 547422	epsilon: 0.1141024600
0848934	steps: 409	lr: 2.560000000000001e-06	evaluation reward:
7.19			
episode: 2158	score: 9.0	memory length: 547893	epsilon: 0.1131698800
0848997	steps: 471	lr: 2.560000000000001e-06	evaluation reward:
7.2			
episode: 2159	score: 4.0	memory length: 548153	epsilon: 0.1126550800
0849032	steps: 260	lr: 2.560000000000001e-06	evaluation reward:
7.17			
episode: 2160	score: 5.0	memory length: 548460	epsilon: 0.1120472200
0849074	steps: 307	lr: 2.560000000000001e-06	evaluation reward:
7.17			
episode: 2161	score: 6.0	memory length: 548797	epsilon: 0.1113799600
084912	steps: 337	lr: 2.560000000000001e-06	evaluation reward: 7
.14			
episode: 2162	score: 7.0	memory length: 549185	epsilon: 0.1106117200
0849172	steps: 388	lr: 2.560000000000001e-06	evaluation reward:
7.17			
episode: 2163	score: 10.0	memory length: 549706	epsilon: 0.109580140
00849242	steps: 521	lr: 2.560000000000001e-06	evaluation reward:
7.24			
episode: 2164	score: 6.0	memory length: 550101	epsilon: 0.1087980400
0849295	steps: 395	lr: 2.560000000000001e-06	evaluation reward:
7.22			
episode: 2165	score: 5.0	memory length: 550442	epsilon: 0.1081228600
0849342	steps: 341	lr: 2.560000000000001e-06	evaluation reward:
7.2			
episode: 2166	score: 11.0	memory length: 551026	epsilon: 0.106966540
0084942	steps: 584	lr: 2.560000000000001e-06	evaluation reward:
7.28			
episode: 2167	score: 10.0	memory length: 551558	epsilon: 0.105913180
00849492	steps: 532	lr: 2.560000000000001e-06	evaluation reward:
7.32			

episode: 2168	score: 8.0	memory length: 551995	epsilon: 0.1050479200
0849551	steps: 437	lr: 2.560000000000001e-06	evaluation reward:
7.36			
episode: 2169	score: 9.0	memory length: 552484	epsilon: 0.1040797000
0849617	steps: 489	lr: 2.560000000000001e-06	evaluation reward:
7.36			
episode: 2170	score: 8.0	memory length: 552764	epsilon: 0.1035253000
0849655	steps: 280	lr: 2.560000000000001e-06	evaluation reward:
7.36			
episode: 2171	score: 10.0	memory length: 553289	epsilon: 0.102485800
00849726	steps: 525	lr: 2.560000000000001e-06	evaluation reward:
7.39			
episode: 2172	score: 9.0	memory length: 553756	epsilon: 0.1015611400
0849789	steps: 467	lr: 2.560000000000001e-06	evaluation reward:
7.42			
episode: 2173	score: 4.0	memory length: 554032	epsilon: 0.1010146600
0849826	steps: 276	lr: 2.560000000000001e-06	evaluation reward:
7.4			
episode: 2174	score: 13.0	memory length: 554656	epsilon: 0.099779140
0084991	steps: 624	lr: 2.560000000000001e-06	evaluation reward:
7.47			
episode: 2175	score: 7.0	memory length: 555061	epsilon: 0.0989772400
0849965	steps: 405	lr: 2.560000000000001e-06	evaluation reward:
7.52			
episode: 2176	score: 22.0	memory length: 555913	epsilon: 0.097290280
0085008	steps: 852	lr: 2.560000000000001e-06	evaluation reward:
7.66			
episode: 2177	score: 7.0	memory length: 556345	epsilon: 0.0964349200
0850139	steps: 432	lr: 2.560000000000001e-06	evaluation reward:
7.65			
episode: 2178	score: 8.0	memory length: 556787	epsilon: 0.0955597600
0850198	steps: 442	lr: 2.560000000000001e-06	evaluation reward:
7.61			
episode: 2179	score: 6.0	memory length: 557139	epsilon: 0.0948628000
0850246	steps: 352	lr: 2.560000000000001e-06	evaluation reward:
7.61			
episode: 2180	score: 5.0	memory length: 557466	epsilon: 0.0942153400
085029	steps: 327	lr: 2.560000000000001e-06	evaluation reward: 7
.59			
episode: 2181	score: 8.0	memory length: 557884	epsilon: 0.0933877000
0850347	steps: 418	lr: 2.560000000000001e-06	evaluation reward:
7.54			
episode: 2182	score: 13.0	memory length: 558535	epsilon: 0.092098720
00850435	steps: 651	lr: 2.560000000000001e-06	evaluation reward:
7.59			
episode: 2183	score: 10.0	memory length: 559045	epsilon: 0.091088920
00850503	steps: 510	lr: 2.560000000000001e-06	evaluation reward:
7.59			
episode: 2184	score: 8.0	memory length: 559466	epsilon: 0.0902553400
085056	steps: 421	lr: 2.560000000000001e-06	evaluation reward: 7
.63			
episode: 2185	score: 11.0	memory length: 559927	epsilon: 0.089342560
00850622	steps: 461	lr: 2.560000000000001e-06	evaluation reward:

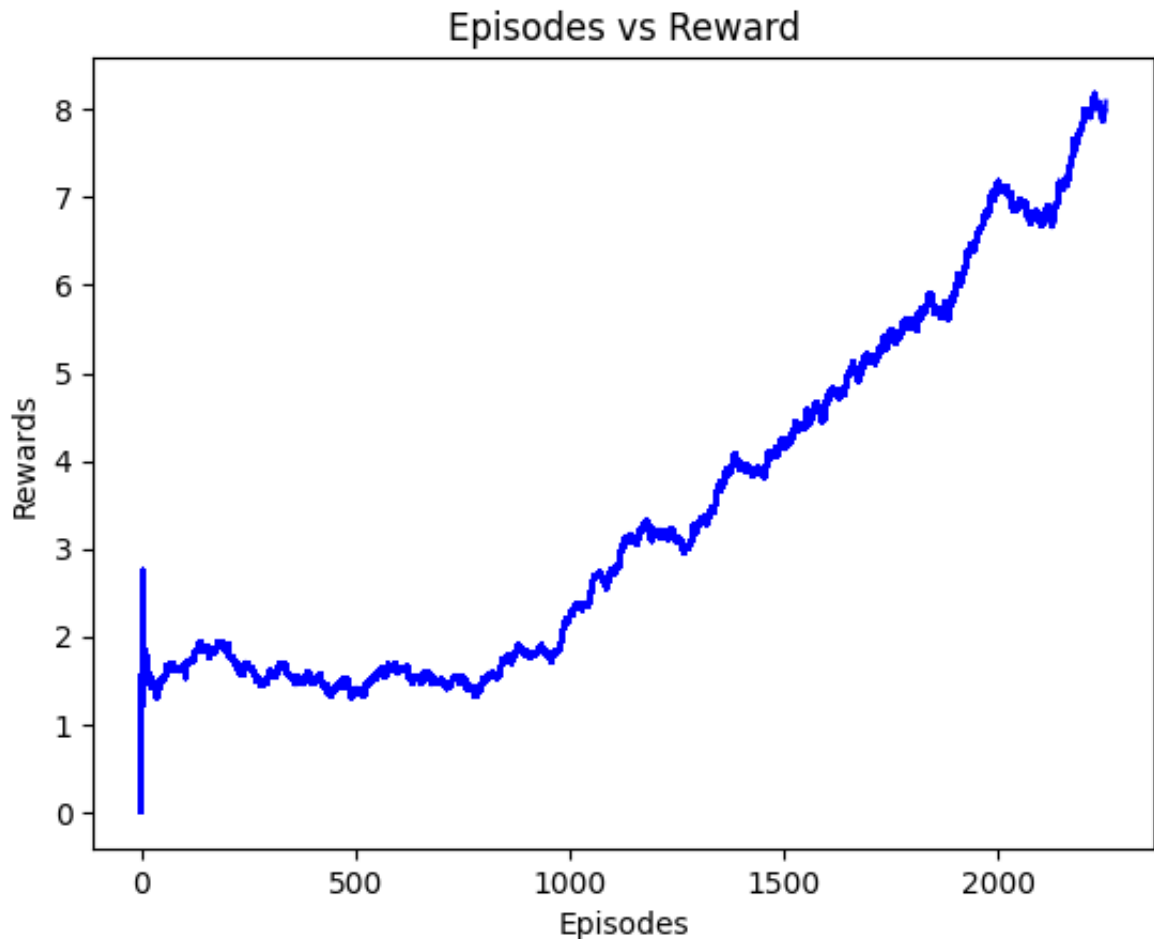
7.62
episode: 2186 score: 8.0 memory length: 560388 epsilon: 0.0884297800
0850685 steps: 461 lr: 2.560000000000001e-06 evaluation reward:
7.64
episode: 2187 score: 9.0 memory length: 560843 epsilon: 0.0875288800
0850746 steps: 455 lr: 2.560000000000001e-06 evaluation reward:
7.65
episode: 2188 score: 10.0 memory length: 561361 epsilon: 0.086503240
00850816 steps: 518 lr: 2.560000000000001e-06 evaluation reward:
7.72
episode: 2189 score: 9.0 memory length: 561828 epsilon: 0.0855785800
0850879 steps: 467 lr: 2.560000000000001e-06 evaluation reward:
7.7
episode: 2190 score: 9.0 memory length: 562339 epsilon: 0.0845668000
0850948 steps: 511 lr: 2.560000000000001e-06 evaluation reward:
7.76
episode: 2191 score: 4.0 memory length: 562633 epsilon: 0.0839846800
0850988 steps: 294 lr: 2.560000000000001e-06 evaluation reward:
7.72
episode: 2192 score: 11.0 memory length: 563157 epsilon: 0.082947160
00851059 steps: 524 lr: 2.560000000000001e-06 evaluation reward:
7.75
episode: 2193 score: 6.0 memory length: 563530 epsilon: 0.0822086200
0851109 steps: 373 lr: 2.560000000000001e-06 evaluation reward:
7.77
episode: 2194 score: 7.0 memory length: 563938 epsilon: 0.0814007800
0851164 steps: 408 lr: 2.560000000000001e-06 evaluation reward:
7.76
episode: 2195 score: 11.0 memory length: 564486 epsilon: 0.080315740
00851238 steps: 548 lr: 2.560000000000001e-06 evaluation reward:
7.81
episode: 2196 score: 7.0 memory length: 564903 epsilon: 0.0794900800
0851294 steps: 417 lr: 2.560000000000001e-06 evaluation reward:
7.8
episode: 2197 score: 9.0 memory length: 565357 epsilon: 0.0785911600
0851356 steps: 454 lr: 2.560000000000001e-06 evaluation reward:
7.81
episode: 2198 score: 9.0 memory length: 565857 epsilon: 0.0776011600
0851423 steps: 500 lr: 2.560000000000001e-06 evaluation reward:
7.84
episode: 2199 score: 7.0 memory length: 566228 epsilon: 0.0768665800
0851473 steps: 371 lr: 2.560000000000001e-06 evaluation reward:
7.85
episode: 2200 score: 11.0 memory length: 566624 epsilon: 0.076082500
00851527 steps: 396 lr: 2.560000000000001e-06 evaluation reward:
7.92
episode: 2201 score: 12.0 memory length: 567208 epsilon: 0.074926180
00851606 steps: 584 lr: 2.560000000000001e-06 evaluation reward:
7.97
episode: 2202 score: 8.0 memory length: 567658 epsilon: 0.0740351800
0851667 steps: 450 lr: 2.560000000000001e-06 evaluation reward:
8.0
episode: 2203 score: 6.0 memory length: 567962 epsilon: 0.0734332600

0851708	steps: 304	lr: 2.5600000000000001e-06	evaluation reward:
7.97			
episode: 2204	score: 4.0	memory length: 568204	epsilon: 0.0729541000
085174	steps: 242	lr: 2.5600000000000001e-06	evaluation reward: 7
.97			
episode: 2205	score: 3.0	memory length: 568434	epsilon: 0.0724987000
0851771	steps: 230	lr: 2.5600000000000001e-06	evaluation reward:
7.91			
episode: 2206	score: 10.0	memory length: 568934	epsilon: 0.071508700
00851839	steps: 500	lr: 2.5600000000000001e-06	evaluation reward:
7.95			
episode: 2207	score: 11.0	memory length: 569505	epsilon: 0.070378120
00851916	steps: 571	lr: 2.5600000000000001e-06	evaluation reward:
8.0			
episode: 2208	score: 10.0	memory length: 570033	epsilon: 0.069332680
00851987	steps: 528	lr: 2.5600000000000001e-06	evaluation reward:
7.97			
episode: 2209	score: 8.0	memory length: 570427	epsilon: 0.0685525600
085204	steps: 394	lr: 2.5600000000000001e-06	evaluation reward: 7
.97			
episode: 2210	score: 10.0	memory length: 570928	epsilon: 0.067560580
00852108	steps: 501	lr: 2.5600000000000001e-06	evaluation reward:
8.0			
episode: 2211	score: 5.0	memory length: 571220	epsilon: 0.0669824200
0852148	steps: 292	lr: 2.5600000000000001e-06	evaluation reward:
7.98			
episode: 2212	score: 9.0	memory length: 571668	epsilon: 0.0660953800
0852208	steps: 448	lr: 2.5600000000000001e-06	evaluation reward:
7.99			
episode: 2213	score: 4.0	memory length: 571910	epsilon: 0.0656162200
0852241	steps: 242	lr: 2.5600000000000001e-06	evaluation reward:
7.98			
episode: 2214	score: 6.0	memory length: 572283	epsilon: 0.0648776800
0852291	steps: 373	lr: 2.5600000000000001e-06	evaluation reward:
7.95			
episode: 2215	score: 8.0	memory length: 572732	epsilon: 0.0639886600
0852352	steps: 449	lr: 2.5600000000000001e-06	evaluation reward:
7.98			
episode: 2216	score: 3.0	memory length: 572945	epsilon: 0.0635669200
085238	steps: 213	lr: 2.5600000000000001e-06	evaluation reward: 7
.91			
episode: 2217	score: 6.0	memory length: 573314	epsilon: 0.0628363000
085243	steps: 369	lr: 2.5600000000000001e-06	evaluation reward: 7
.93			
episode: 2218	score: 9.0	memory length: 573785	epsilon: 0.0619037200
0852494	steps: 471	lr: 2.5600000000000001e-06	evaluation reward:
7.96			
episode: 2219	score: 7.0	memory length: 574170	epsilon: 0.0611414200
0852546	steps: 385	lr: 2.5600000000000001e-06	evaluation reward:
7.98			
episode: 2220	score: 12.0	memory length: 574799	epsilon: 0.059896000
00852631	steps: 629	lr: 2.5600000000000001e-06	evaluation reward:
8.05			

episode: 2221	score: 7.0	memory length: 575209	epsilon: 0.0590842000
0852686	steps: 410	lr: 2.560000000000001e-06	evaluation reward:
8.07			
episode: 2222	score: 10.0	memory length: 575721	epsilon: 0.058070440
008527555	steps: 512	lr: 2.560000000000001e-06	evaluation reward
: 8.1			
episode: 2223	score: 8.0	memory length: 576159	epsilon: 0.0572032000
08528146	steps: 438	lr: 2.560000000000001e-06	evaluation reward:
8.14			
episode: 2224	score: 5.0	memory length: 576506	epsilon: 0.0565161400
08528615	steps: 347	lr: 2.560000000000001e-06	evaluation reward:
8.12			
episode: 2225	score: 8.0	memory length: 576931	epsilon: 0.0556746400
0852919	steps: 425	lr: 2.560000000000001e-06	evaluation reward:
8.16			
episode: 2226	score: 10.0	memory length: 577451	epsilon: 0.054645040
00852989	steps: 520	lr: 2.560000000000001e-06	evaluation reward:
8.18			
episode: 2227	score: 6.0	memory length: 577827	epsilon: 0.0539005600
085304	steps: 376	lr: 2.560000000000001e-06	evaluation reward: 8
.14			
episode: 2228	score: 6.0	memory length: 578164	epsilon: 0.0532333000
08530854	steps: 337	lr: 2.560000000000001e-06	evaluation reward:
8.07			
episode: 2229	score: 8.0	memory length: 578579	epsilon: 0.0524116000
08531414	steps: 415	lr: 2.560000000000001e-06	evaluation reward:
8.11			
episode: 2230	score: 5.0	memory length: 578888	epsilon: 0.0517997800
0853183	steps: 309	lr: 2.560000000000001e-06	evaluation reward:
8.04			
episode: 2231	score: 7.0	memory length: 579313	epsilon: 0.0509582800
08532406	steps: 425	lr: 2.560000000000001e-06	evaluation reward:
8.05			
episode: 2232	score: 7.0	memory length: 579693	epsilon: 0.0502058800
0853292	steps: 380	lr: 2.560000000000001e-06	evaluation reward:
8.03			
episode: 2233	score: 9.0	memory length: 580167	epsilon: 0.0492673600
0853356	steps: 474	lr: 2.560000000000001e-06	evaluation reward:
8.06			
episode: 2234	score: 8.0	memory length: 580619	epsilon: 0.0483724000
0853417	steps: 452	lr: 2.560000000000001e-06	evaluation reward:
8.01			
episode: 2235	score: 10.0	memory length: 581060	epsilon: 0.047499220
008534765	steps: 441	lr: 2.560000000000001e-06	evaluation reward
: 8.05			
episode: 2236	score: 4.0	memory length: 581335	epsilon: 0.0469547200
08535136	steps: 275	lr: 2.560000000000001e-06	evaluation reward:
8.02			
episode: 2237	score: 6.0	memory length: 581726	epsilon: 0.0461805400
08535664	steps: 391	lr: 2.560000000000001e-06	evaluation reward:
8.0			
episode: 2238	score: 15.0	memory length: 582316	epsilon: 0.045012340
00853646	steps: 590	lr: 2.560000000000001e-06	evaluation reward:

8.07
episode: 2239 score: 6.0 memory length: 582680 epsilon: 0.0442916200
0853695 steps: 364 lr: 2.560000000000001e-06 evaluation reward:
8.05
episode: 2240 score: 8.0 memory length: 583134 epsilon: 0.0433927000
08537566 steps: 454 lr: 2.560000000000001e-06 evaluation reward:
8.04
episode: 2241 score: 7.0 memory length: 583536 epsilon: 0.0425967400
0853811 steps: 402 lr: 2.560000000000001e-06 evaluation reward:
8.02
episode: 2242 score: 8.0 memory length: 583990 epsilon: 0.0416978200
0853872 steps: 454 lr: 2.560000000000001e-06 evaluation reward:
7.99
episode: 2243 score: 4.0 memory length: 584283 epsilon: 0.0411176800
0853912 steps: 293 lr: 2.560000000000001e-06 evaluation reward:
7.91
episode: 2244 score: 12.0 memory length: 584742 epsilon: 0.040208860
0853974 steps: 459 lr: 2.560000000000001e-06 evaluation reward:
7.88
episode: 2245 score: 4.0 memory length: 585004 epsilon: 0.0396901000
0854009 steps: 262 lr: 2.560000000000001e-06 evaluation reward:
7.87
episode: 2246 score: 8.0 memory length: 585405 epsilon: 0.0388961200
0854063 steps: 401 lr: 2.560000000000001e-06 evaluation reward:
7.89
episode: 2247 score: 6.0 memory length: 585740 epsilon: 0.0382328200
08541085 steps: 335 lr: 2.560000000000001e-06 evaluation reward:
7.9
episode: 2248 score: 8.0 memory length: 586211 epsilon: 0.0373002400
0854172 steps: 471 lr: 2.560000000000001e-06 evaluation reward:
7.88
episode: 2249 score: 14.0 memory length: 586773 epsilon: 0.036187480
0854248 steps: 562 lr: 2.560000000000001e-06 evaluation reward:
7.95
episode: 2250 score: 10.0 memory length: 587265 epsilon: 0.035213320
08543145 steps: 492 lr: 2.560000000000001e-06 evaluation reward
: 7.98
episode: 2251 score: 7.0 memory length: 587669 epsilon: 0.0344134000
0854369 steps: 404 lr: 2.560000000000001e-06 evaluation reward:
7.99
episode: 2252 score: 14.0 memory length: 588209 epsilon: 0.033344200
0854442 steps: 540 lr: 2.560000000000001e-06 evaluation reward:
8.09

```
-----  
KeyboardInterrupt                                Traceback (most recent call last)  
)  
Cell In[5], line 41  
    39 # Start training after random sample generation  
    40 if(frame >= train_frame): # You can set train_frame to a lower value while testing your starts training earlier  
----> 41     agent.train_policy_net(frame)  
    42     # Update the target network only for Double DQN only  
    43     if double_dqn and (frame % update_target_network_frequency)==  
0:  
  
File ~/assignment5/agent.py:66, in Agent.train_policy_net(self, frame)  
    63 if self.epsilon > self.epsilon_min:  
    64     self.epsilon -= self.epsilon_decay  
----> 66 mini_batch = self.memory.sample_mini_batch(frame)  
    67 mini_batch = np.array(mini_batch, dtype=object).transpose()  
    69 history = np.stack(mini_batch[0], axis=0)  
  
File ~/assignment5/memory.py:28, in ReplayMemory.sample_mini_batch(self, frame)  
    26 sample = []  
    27 for j in range(HISTORY_SIZE + 1):  
----> 28     sample.append(self.memory[i + j])  
    30 sample = np.array(sample, dtype=object)  
    31 mini_batch.append((np.stack(sample[:, 0], axis=0), sample[3, 1], sample[3, 2], sample[3, 3]))  
  
KeyboardInterrupt:
```



Trianing a double DQN

```
In [ ]: device = torch.device("cuda:0" if torch.cuda.is_available() else "cpu")

rewards, episodes = [], []
best_eval_reward = 0
for e in range(EPIISODES):
    done = False
    score = 0

    history = np.zeros([5, 84, 84], dtype=np.uint8)
    step = 0
    state, _ = env.reset()
    next_state = state
    life = number_lives

    get_init_state(history, state, HISTORY_SIZE)

    while not done:
        step += 1
        frame += 1

        # Perform a fire action if ball is no longer on screen to continu
        if step > 1 and len(np.unique(next_state[:189] == state[:189])) <
```

```

        action = torch.tensor([[0]]).cuda()
    else:
        action = agent.get_action(np.float32(history[:4, :, :]) / 255)
    state = next_state
    next_state, reward, terminated, truncated, info = env.step(action)
    done = terminated or truncated

    frame_next_state = get_frame(next_state)
    history[4, :, :] = frame_next_state
    terminal_state = check_live(life, info['lives'])

    life = info['lives']
    r = reward

    # Store the transition in memory
    agent.memory.push(deepcopy(frame_next_state), action.cpu(), r, te
    # Start training after random sample generation
    if(frame >= train_frame): # You can set train_frame to a lower va
        agent.train_policy_net(frame)
        # Update the target network only for Double DQN only
        if double_dqn and (frame % update_target_network_frequency)==
            agent.update_target_net()
    score += reward
    history[:4, :, :] = history[1:, :, :]

    if done:
        evaluation_reward.append(score)
        rewards.append(np.mean(evaluation_reward))
        episodes.append(e)
        pylab.plot(episodes, rewards, 'b')
        pylab.xlabel('Episodes')
        pylab.ylabel('Rewards')
        pylab.title('Episodes vs Reward')
        pylab.savefig("./save_graph/breakout_dqn.png") # save graph f

    # every episode, plot the play time
    print("episode:", e, " score:", score, " memory length:",
          len(agent.memory), " epsilon:", agent.epsilon, " ste
          " lr:", agent.optimizer.param_groups[0]['lr'], " e

    # if the mean of scores of last 100 episode is bigger than 5
    ### Change this save condition to whatever you prefer ###
    if np.mean(evaluation_reward) > 6 and np.mean(evaluation_rewa
        save_dir = "./save_model/breakout_ddqn_"+str(e)+".pth"
        torch.save(agent.policy_net, save_dir)
        best_eval_reward = np.mean(evaluation_reward)

```

```

episode: 0    score: 3.0    memory length: 249    epsilon: 1.0    steps: 249
lr: 0.0001    evaluation reward: 3.0
episode: 1    score: 3.0    memory length: 497    epsilon: 1.0    steps: 248
lr: 0.0001    evaluation reward: 3.0
episode: 2    score: 2.0    memory length: 694    epsilon: 1.0    steps: 197
lr: 0.0001    evaluation reward: 2.6666666666666665
episode: 3    score: 2.0    memory length: 891    epsilon: 1.0    steps: 197

```

```
lr: 0.0001      evaluation reward: 2.5
episode: 4      score: 3.0    memory length: 1138    epsilon: 1.0    steps: 247
lr: 0.0001      evaluation reward: 2.6
episode: 5      score: 0.0    memory length: 1260    epsilon: 1.0    steps: 122
lr: 0.0001      evaluation reward: 2.1666666666666665
episode: 6      score: 4.0    memory length: 1517    epsilon: 1.0    steps: 257
lr: 0.0001      evaluation reward: 2.4285714285714284
episode: 7      score: 1.0    memory length: 1685    epsilon: 1.0    steps: 168
lr: 0.0001      evaluation reward: 2.25
episode: 8      score: 0.0    memory length: 1808    epsilon: 1.0    steps: 123
lr: 0.0001      evaluation reward: 2.0
episode: 9      score: 3.0    memory length: 2018    epsilon: 1.0    steps: 210
lr: 0.0001      evaluation reward: 2.1
episode: 10     score: 0.0    memory length: 2141    epsilon: 1.0    steps: 12
3  lr: 0.0001      evaluation reward: 1.9090909090909092
episode: 11     score: 0.0    memory length: 2263    epsilon: 1.0    steps: 12
2  lr: 0.0001      evaluation reward: 1.75
episode: 12     score: 2.0    memory length: 2479    epsilon: 1.0    steps: 21
6  lr: 0.0001      evaluation reward: 1.7692307692307692
episode: 13     score: 0.0    memory length: 2601    epsilon: 1.0    steps: 12
2  lr: 0.0001      evaluation reward: 1.6428571428571428
episode: 14     score: 4.0    memory length: 2879    epsilon: 1.0    steps: 27
8  lr: 0.0001      evaluation reward: 1.8
episode: 15     score: 2.0    memory length: 3077    epsilon: 1.0    steps: 19
8  lr: 0.0001      evaluation reward: 1.8125
episode: 16     score: 3.0    memory length: 3306    epsilon: 1.0    steps: 22
9  lr: 0.0001      evaluation reward: 1.8823529411764706
episode: 17     score: 0.0    memory length: 3429    epsilon: 1.0    steps: 12
3  lr: 0.0001      evaluation reward: 1.7777777777777777
episode: 18     score: 6.0    memory length: 3804    epsilon: 1.0    steps: 37
5  lr: 0.0001      evaluation reward: 2.0
episode: 19     score: 4.0    memory length: 4080    epsilon: 1.0    steps: 27
6  lr: 0.0001      evaluation reward: 2.1
episode: 20     score: 3.0    memory length: 4327    epsilon: 1.0    steps: 24
7  lr: 0.0001      evaluation reward: 2.142857142857143
episode: 21     score: 0.0    memory length: 4449    epsilon: 1.0    steps: 12
2  lr: 0.0001      evaluation reward: 2.0454545454545454
episode: 22     score: 3.0    memory length: 4678    epsilon: 1.0    steps: 22
9  lr: 0.0001      evaluation reward: 2.0869565217391304
episode: 23     score: 0.0    memory length: 4801    epsilon: 1.0    steps: 12
3  lr: 0.0001      evaluation reward: 2.0
episode: 24     score: 3.0    memory length: 5046    epsilon: 1.0    steps: 24
5  lr: 0.0001      evaluation reward: 2.04
episode: 25     score: 3.0    memory length: 5275    epsilon: 1.0    steps: 22
9  lr: 0.0001      evaluation reward: 2.076923076923077
episode: 26     score: 0.0    memory length: 5397    epsilon: 1.0    steps: 12
2  lr: 0.0001      evaluation reward: 2.0
episode: 27     score: 1.0    memory length: 5566    epsilon: 1.0    steps: 16
9  lr: 0.0001      evaluation reward: 1.9642857142857142
episode: 28     score: 0.0    memory length: 5689    epsilon: 1.0    steps: 12
3  lr: 0.0001      evaluation reward: 1.896551724137931
episode: 29     score: 4.0    memory length: 5988    epsilon: 1.0    steps: 29
9  lr: 0.0001      evaluation reward: 1.9666666666666666
```

```
episode: 30  score: 0.0  memory length: 6111  epsilon: 1.0  steps: 12
3    lr: 0.0001  evaluation reward: 1.903225806451613
episode: 31  score: 0.0  memory length: 6234  epsilon: 1.0  steps: 12
3    lr: 0.0001  evaluation reward: 1.84375
episode: 32  score: 0.0  memory length: 6357  epsilon: 1.0  steps: 12
3    lr: 0.0001  evaluation reward: 1.7878787878787878
episode: 33  score: 0.0  memory length: 6479  epsilon: 1.0  steps: 12
2    lr: 0.0001  evaluation reward: 1.7352941176470589
episode: 34  score: 2.0  memory length: 6694  epsilon: 1.0  steps: 21
5    lr: 0.0001  evaluation reward: 1.7428571428571429
episode: 35  score: 0.0  memory length: 6816  epsilon: 1.0  steps: 12
2    lr: 0.0001  evaluation reward: 1.6944444444444444
episode: 36  score: 2.0  memory length: 7014  epsilon: 1.0  steps: 19
8    lr: 0.0001  evaluation reward: 1.7027027027027026
episode: 37  score: 2.0  memory length: 7212  epsilon: 1.0  steps: 19
8    lr: 0.0001  evaluation reward: 1.7105263157894737
episode: 38  score: 6.0  memory length: 7548  epsilon: 1.0  steps: 33
6    lr: 0.0001  evaluation reward: 1.8205128205128205
episode: 39  score: 2.0  memory length: 7748  epsilon: 1.0  steps: 20
0    lr: 0.0001  evaluation reward: 1.825
episode: 40  score: 2.0  memory length: 7965  epsilon: 1.0  steps: 21
7    lr: 0.0001  evaluation reward: 1.829268292682927
episode: 41  score: 2.0  memory length: 8164  epsilon: 1.0  steps: 19
9    lr: 0.0001  evaluation reward: 1.8333333333333333
episode: 42  score: 1.0  memory length: 8333  epsilon: 1.0  steps: 16
9    lr: 0.0001  evaluation reward: 1.813953488372093
episode: 43  score: 1.0  memory length: 8501  epsilon: 1.0  steps: 16
8    lr: 0.0001  evaluation reward: 1.7954545454545454
episode: 44  score: 1.0  memory length: 8670  epsilon: 1.0  steps: 16
9    lr: 0.0001  evaluation reward: 1.7777777777777777
episode: 45  score: 4.0  memory length: 8991  epsilon: 1.0  steps: 32
1    lr: 0.0001  evaluation reward: 1.826086956521739
episode: 46  score: 1.0  memory length: 9142  epsilon: 1.0  steps: 15
1    lr: 0.0001  evaluation reward: 1.8085106382978724
episode: 47  score: 1.0  memory length: 9312  epsilon: 1.0  steps: 17
0    lr: 0.0001  evaluation reward: 1.7916666666666667
episode: 48  score: 3.0  memory length: 9540  epsilon: 1.0  steps: 22
8    lr: 0.0001  evaluation reward: 1.816326530612245
episode: 49  score: 3.0  memory length: 9788  epsilon: 1.0  steps: 24
8    lr: 0.0001  evaluation reward: 1.84
episode: 50  score: 2.0  memory length: 9986  epsilon: 1.0  steps: 19
8    lr: 0.0001  evaluation reward: 1.8431372549019607
episode: 51  score: 3.0  memory length: 10251  epsilon: 1.0  steps: 2
65   lr: 0.0001  evaluation reward: 1.8653846153846154
episode: 52  score: 3.0  memory length: 10477  epsilon: 1.0  steps: 2
26   lr: 0.0001  evaluation reward: 1.8867924528301887
episode: 53  score: 4.0  memory length: 10743  epsilon: 1.0  steps: 2
66   lr: 0.0001  evaluation reward: 1.9259259259259258
episode: 54  score: 0.0  memory length: 10866  epsilon: 1.0  steps: 1
23   lr: 0.0001  evaluation reward: 1.8909090909090909
episode: 55  score: 2.0  memory length: 11048  epsilon: 1.0  steps: 1
82   lr: 0.0001  evaluation reward: 1.8928571428571428
episode: 56  score: 2.0  memory length: 11245  epsilon: 1.0  steps: 1
```

```
97   lr: 0.0001      evaluation reward: 1.894736842105263
episode: 57   score: 3.0   memory length: 11512   epsilon: 1.0   steps: 2
67   lr: 0.0001      evaluation reward: 1.9137931034482758
episode: 58   score: 1.0   memory length: 11662   epsilon: 1.0   steps: 1
50   lr: 0.0001      evaluation reward: 1.8983050847457628
episode: 59   score: 4.0   memory length: 11940   epsilon: 1.0   steps: 2
78   lr: 0.0001      evaluation reward: 1.9333333333333333
episode: 60   score: 1.0   memory length: 12091   epsilon: 1.0   steps: 1
51   lr: 0.0001      evaluation reward: 1.9180327868852458
episode: 61   score: 3.0   memory length: 12301   epsilon: 1.0   steps: 2
10   lr: 0.0001      evaluation reward: 1.935483870967742
episode: 62   score: 0.0   memory length: 12424   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.9047619047619047
episode: 63   score: 3.0   memory length: 12671   epsilon: 1.0   steps: 2
47   lr: 0.0001      evaluation reward: 1.921875
episode: 64   score: 1.0   memory length: 12839   epsilon: 1.0   steps: 1
68   lr: 0.0001      evaluation reward: 1.9076923076923078
episode: 65   score: 5.0   memory length: 13181   epsilon: 1.0   steps: 3
42   lr: 0.0001      evaluation reward: 1.9545454545454546
episode: 66   score: 5.0   memory length: 13491   epsilon: 1.0   steps: 3
10   lr: 0.0001      evaluation reward: 2.0
episode: 67   score: 0.0   memory length: 13613   epsilon: 1.0   steps: 1
22   lr: 0.0001      evaluation reward: 1.9705882352941178
episode: 68   score: 0.0   memory length: 13736   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.9420289855072463
episode: 69   score: 0.0   memory length: 13859   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.9142857142857144
episode: 70   score: 2.0   memory length: 14056   epsilon: 1.0   steps: 1
97   lr: 0.0001      evaluation reward: 1.9154929577464788
episode: 71   score: 0.0   memory length: 14179   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.8888888888888888
episode: 72   score: 2.0   memory length: 14376   epsilon: 1.0   steps: 1
97   lr: 0.0001      evaluation reward: 1.8904109589041096
episode: 73   score: 0.0   memory length: 14499   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.864864864864865
episode: 74   score: 1.0   memory length: 14650   epsilon: 1.0   steps: 1
51   lr: 0.0001      evaluation reward: 1.8533333333333333
episode: 75   score: 1.0   memory length: 14819   epsilon: 1.0   steps: 1
69   lr: 0.0001      evaluation reward: 1.8421052631578947
episode: 76   score: 3.0   memory length: 15066   epsilon: 1.0   steps: 2
47   lr: 0.0001      evaluation reward: 1.8571428571428572
episode: 77   score: 3.0   memory length: 15294   epsilon: 1.0   steps: 2
28   lr: 0.0001      evaluation reward: 1.8717948717948718
episode: 78   score: 0.0   memory length: 15417   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.8481012658227849
episode: 79   score: 2.0   memory length: 15615   epsilon: 1.0   steps: 1
98   lr: 0.0001      evaluation reward: 1.85
episode: 80   score: 0.0   memory length: 15738   epsilon: 1.0   steps: 1
23   lr: 0.0001      evaluation reward: 1.8271604938271604
episode: 81   score: 1.0   memory length: 15888   epsilon: 1.0   steps: 1
50   lr: 0.0001      evaluation reward: 1.8170731707317074
episode: 82   score: 1.0   memory length: 16038   epsilon: 1.0   steps: 1
50   lr: 0.0001      evaluation reward: 1.8072289156626506
```



```
episode: 83  score: 1.0  memory length: 16207  epsilon: 1.0  steps: 1
69  lr: 0.0001  evaluation reward: 1.7976190476190477
episode: 84  score: 1.0  memory length: 16376  epsilon: 1.0  steps: 1
69  lr: 0.0001  evaluation reward: 1.7882352941176471
episode: 85  score: 3.0  memory length: 16622  epsilon: 1.0  steps: 2
46  lr: 0.0001  evaluation reward: 1.802325581395349
episode: 86  score: 1.0  memory length: 16772  epsilon: 1.0  steps: 1
50  lr: 0.0001  evaluation reward: 1.793103448275862
episode: 87  score: 0.0  memory length: 16895  epsilon: 1.0  steps: 1
23  lr: 0.0001  evaluation reward: 1.7727272727272727
episode: 88  score: 0.0  memory length: 17018  epsilon: 1.0  steps: 1
23  lr: 0.0001  evaluation reward: 1.752808988764045
episode: 89  score: 4.0  memory length: 17312  epsilon: 1.0  steps: 2
94  lr: 0.0001  evaluation reward: 1.7777777777777777
episode: 90  score: 2.0  memory length: 17510  epsilon: 1.0  steps: 1
98  lr: 0.0001  evaluation reward: 1.7802197802197801
episode: 91  score: 0.0  memory length: 17633  epsilon: 1.0  steps: 1
23  lr: 0.0001  evaluation reward: 1.7608695652173914
episode: 92  score: 2.0  memory length: 17830  epsilon: 1.0  steps: 1
97  lr: 0.0001  evaluation reward: 1.7634408602150538
episode: 93  score: 0.0  memory length: 17953  epsilon: 1.0  steps: 1
23  lr: 0.0001  evaluation reward: 1.7446808510638299
episode: 94  score: 3.0  memory length: 18198  epsilon: 1.0  steps: 2
45  lr: 0.0001  evaluation reward: 1.7578947368421052
episode: 95  score: 1.0  memory length: 18367  epsilon: 1.0  steps: 1
69  lr: 0.0001  evaluation reward: 1.75
episode: 96  score: 2.0  memory length: 18565  epsilon: 1.0  steps: 1
98  lr: 0.0001  evaluation reward: 1.7525773195876289
episode: 97  score: 4.0  memory length: 18841  epsilon: 1.0  steps: 2
76  lr: 0.0001  evaluation reward: 1.7755102040816326
episode: 98  score: 3.0  memory length: 19068  epsilon: 1.0  steps: 2
27  lr: 0.0001  evaluation reward: 1.7878787878787878
episode: 99  score: 1.0  memory length: 19236  epsilon: 1.0  steps: 1
68  lr: 0.0001  evaluation reward: 1.78
episode: 100  score: 1.0  memory length: 19387  epsilon: 1.0  steps:
151  lr: 0.0001  evaluation reward: 1.76
episode: 101  score: 1.0  memory length: 19537  epsilon: 1.0  steps:
150  lr: 0.0001  evaluation reward: 1.74
episode: 102  score: 3.0  memory length: 19784  epsilon: 1.0  steps:
247  lr: 0.0001  evaluation reward: 1.75
episode: 103  score: 1.0  memory length: 19935  epsilon: 1.0  steps:
151  lr: 0.0001  evaluation reward: 1.74
episode: 104  score: 2.0  memory length: 20136  epsilon: 1.0  steps:
201  lr: 0.0001  evaluation reward: 1.73
episode: 105  score: 2.0  memory length: 20334  epsilon: 1.0  steps:
198  lr: 0.0001  evaluation reward: 1.75
episode: 106  score: 1.0  memory length: 20505  epsilon: 1.0  steps:
171  lr: 0.0001  evaluation reward: 1.72
episode: 107  score: 3.0  memory length: 20734  epsilon: 1.0  steps:
229  lr: 0.0001  evaluation reward: 1.74
episode: 108  score: 1.0  memory length: 20884  epsilon: 1.0  steps:
150  lr: 0.0001  evaluation reward: 1.75
episode: 109  score: 0.0  memory length: 21006  epsilon: 1.0  steps:
```

122	lr: 0.0001	evaluation reward: 1.72		
episode: 110	score: 0.0	memory length: 21129	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.72		
episode: 111	score: 2.0	memory length: 21327	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.74		
episode: 112	score: 2.0	memory length: 21548	epsilon: 1.0	steps:
221	lr: 0.0001	evaluation reward: 1.74		
episode: 113	score: 2.0	memory length: 21745	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.76		
episode: 114	score: 1.0	memory length: 21916	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.73		
episode: 115	score: 2.0	memory length: 22134	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.73		
episode: 116	score: 4.0	memory length: 22410	epsilon: 1.0	steps:
276	lr: 0.0001	evaluation reward: 1.74		
episode: 117	score: 3.0	memory length: 22657	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.77		
episode: 118	score: 0.0	memory length: 22779	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.71		
episode: 119	score: 1.0	memory length: 22929	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.68		
episode: 120	score: 1.0	memory length: 23079	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.66		
episode: 121	score: 4.0	memory length: 23377	epsilon: 1.0	steps:
298	lr: 0.0001	evaluation reward: 1.7		
episode: 122	score: 2.0	memory length: 23579	epsilon: 1.0	steps:
202	lr: 0.0001	evaluation reward: 1.69		
episode: 123	score: 1.0	memory length: 23750	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.7		
episode: 124	score: 0.0	memory length: 23872	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.67		
episode: 125	score: 0.0	memory length: 23995	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.64		
episode: 126	score: 2.0	memory length: 24193	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.66		
episode: 127	score: 3.0	memory length: 24423	epsilon: 1.0	steps:
230	lr: 0.0001	evaluation reward: 1.68		
episode: 128	score: 1.0	memory length: 24593	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.69		
episode: 129	score: 1.0	memory length: 24744	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.66		
episode: 130	score: 1.0	memory length: 24915	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.67		
episode: 131	score: 0.0	memory length: 25038	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.67		
episode: 132	score: 1.0	memory length: 25209	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.68		
episode: 133	score: 2.0	memory length: 25426	epsilon: 1.0	steps:
217	lr: 0.0001	evaluation reward: 1.7		
episode: 134	score: 1.0	memory length: 25595	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.69		
episode: 135	score: 2.0	memory length: 25796	epsilon: 1.0	steps:
201	lr: 0.0001	evaluation reward: 1.71		

episode: 136	score: 2.0	memory length: 26012	epsilon: 1.0	steps:
216	lr: 0.0001	evaluation reward: 1.71		
episode: 137	score: 2.0	memory length: 26230	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.71		
episode: 138	score: 2.0	memory length: 26448	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.67		
episode: 139	score: 2.0	memory length: 26645	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.67		
episode: 140	score: 1.0	memory length: 26813	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.66		
episode: 141	score: 2.0	memory length: 27030	epsilon: 1.0	steps:
217	lr: 0.0001	evaluation reward: 1.66		
episode: 142	score: 0.0	memory length: 27152	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.65		
episode: 143	score: 1.0	memory length: 27323	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.65		
episode: 144	score: 2.0	memory length: 27521	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.66		
episode: 145	score: 2.0	memory length: 27737	epsilon: 1.0	steps:
216	lr: 0.0001	evaluation reward: 1.64		
episode: 146	score: 1.0	memory length: 27888	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.64		
episode: 147	score: 1.0	memory length: 28039	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.64		
episode: 148	score: 3.0	memory length: 28265	epsilon: 1.0	steps:
226	lr: 0.0001	evaluation reward: 1.64		
episode: 149	score: 1.0	memory length: 28437	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.62		
episode: 150	score: 3.0	memory length: 28687	epsilon: 1.0	steps:
250	lr: 0.0001	evaluation reward: 1.63		
episode: 151	score: 1.0	memory length: 28837	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.61		
episode: 152	score: 1.0	memory length: 29006	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.59		
episode: 153	score: 0.0	memory length: 29129	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.55		
episode: 154	score: 4.0	memory length: 29405	epsilon: 1.0	steps:
276	lr: 0.0001	evaluation reward: 1.59		
episode: 155	score: 0.0	memory length: 29528	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.57		
episode: 156	score: 1.0	memory length: 29679	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.56		
episode: 157	score: 0.0	memory length: 29802	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.53		
episode: 158	score: 0.0	memory length: 29924	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.52		
episode: 159	score: 2.0	memory length: 30142	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.5		
episode: 160	score: 0.0	memory length: 30265	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.49		
episode: 161	score: 2.0	memory length: 30480	epsilon: 1.0	steps:
215	lr: 0.0001	evaluation reward: 1.48		
episode: 162	score: 4.0	memory length: 30802	epsilon: 1.0	steps:

322	lr: 0.0001	evaluation reward: 1.52		
episode: 163	score: 0.0	memory length: 30924	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.49		
episode: 164	score: 2.0	memory length: 31143	epsilon: 1.0	steps:
219	lr: 0.0001	evaluation reward: 1.5		
episode: 165	score: 1.0	memory length: 31312	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.46		
episode: 166	score: 2.0	memory length: 31529	epsilon: 1.0	steps:
217	lr: 0.0001	evaluation reward: 1.43		
episode: 167	score: 2.0	memory length: 31746	epsilon: 1.0	steps:
217	lr: 0.0001	evaluation reward: 1.45		
episode: 168	score: 1.0	memory length: 31917	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.46		
episode: 169	score: 6.0	memory length: 32169	epsilon: 1.0	steps:
252	lr: 0.0001	evaluation reward: 1.52		
episode: 170	score: 0.0	memory length: 32292	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.5		
episode: 171	score: 2.0	memory length: 32508	epsilon: 1.0	steps:
216	lr: 0.0001	evaluation reward: 1.52		
episode: 172	score: 0.0	memory length: 32630	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.5		
episode: 173	score: 4.0	memory length: 32920	epsilon: 1.0	steps:
290	lr: 0.0001	evaluation reward: 1.54		
episode: 174	score: 2.0	memory length: 33136	epsilon: 1.0	steps:
216	lr: 0.0001	evaluation reward: 1.55		
episode: 175	score: 1.0	memory length: 33308	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.55		
episode: 176	score: 0.0	memory length: 33430	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.52		
episode: 177	score: 2.0	memory length: 33612	epsilon: 1.0	steps:
182	lr: 0.0001	evaluation reward: 1.51		
episode: 178	score: 2.0	memory length: 33828	epsilon: 1.0	steps:
216	lr: 0.0001	evaluation reward: 1.53		
episode: 179	score: 0.0	memory length: 33950	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.51		
episode: 180	score: 2.0	memory length: 34129	epsilon: 1.0	steps:
179	lr: 0.0001	evaluation reward: 1.53		
episode: 181	score: 1.0	memory length: 34301	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.53		
episode: 182	score: 0.0	memory length: 34424	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.52		
episode: 183	score: 3.0	memory length: 34688	epsilon: 1.0	steps:
264	lr: 0.0001	evaluation reward: 1.54		
episode: 184	score: 4.0	memory length: 34965	epsilon: 1.0	steps:
277	lr: 0.0001	evaluation reward: 1.57		
episode: 185	score: 0.0	memory length: 35088	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.54		
episode: 186	score: 0.0	memory length: 35210	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.53		
episode: 187	score: 4.0	memory length: 35504	epsilon: 1.0	steps:
294	lr: 0.0001	evaluation reward: 1.57		
episode: 188	score: 1.0	memory length: 35675	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.58		

episode: 189	score: 0.0	memory length: 35798	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.54		
episode: 190	score: 7.0	memory length: 36217	epsilon: 1.0	steps:
419	lr: 0.0001	evaluation reward: 1.59		
episode: 191	score: 1.0	memory length: 36385	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.6		
episode: 192	score: 0.0	memory length: 36508	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.58		
episode: 193	score: 3.0	memory length: 36775	epsilon: 1.0	steps:
267	lr: 0.0001	evaluation reward: 1.61		
episode: 194	score: 1.0	memory length: 36944	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.59		
episode: 195	score: 0.0	memory length: 37067	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.58		
episode: 196	score: 1.0	memory length: 37218	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.57		
episode: 197	score: 2.0	memory length: 37416	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.55		
episode: 198	score: 0.0	memory length: 37539	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.52		
episode: 199	score: 2.0	memory length: 37758	epsilon: 1.0	steps:
219	lr: 0.0001	evaluation reward: 1.53		
episode: 200	score: 1.0	memory length: 37927	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.53		
episode: 201	score: 0.0	memory length: 38049	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.52		
episode: 202	score: 0.0	memory length: 38172	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.49		
episode: 203	score: 2.0	memory length: 38369	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.5		
episode: 204	score: 2.0	memory length: 38567	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.5		
episode: 205	score: 4.0	memory length: 38835	epsilon: 1.0	steps:
268	lr: 0.0001	evaluation reward: 1.52		
episode: 206	score: 3.0	memory length: 39046	epsilon: 1.0	steps:
211	lr: 0.0001	evaluation reward: 1.54		
episode: 207	score: 2.0	memory length: 39264	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.53		
episode: 208	score: 2.0	memory length: 39462	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.54		
episode: 209	score: 3.0	memory length: 39709	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.57		
episode: 210	score: 3.0	memory length: 39976	epsilon: 1.0	steps:
267	lr: 0.0001	evaluation reward: 1.6		
episode: 211	score: 3.0	memory length: 40222	epsilon: 1.0	steps:
246	lr: 0.0001	evaluation reward: 1.61		
episode: 212	score: 2.0	memory length: 40420	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.61		
episode: 213	score: 1.0	memory length: 40590	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.6		
episode: 214	score: 2.0	memory length: 40788	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.61		
episode: 215	score: 1.0	memory length: 40957	epsilon: 1.0	steps:

169	lr: 0.0001	evaluation reward: 1.6		
episode: 216	score: 3.0	memory length: 41203	epsilon: 1.0	steps:
246	lr: 0.0001	evaluation reward: 1.59		
episode: 217	score: 0.0	memory length: 41325	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.56		
episode: 218	score: 1.0	memory length: 41477	epsilon: 1.0	steps:
152	lr: 0.0001	evaluation reward: 1.57		
episode: 219	score: 4.0	memory length: 41771	epsilon: 1.0	steps:
294	lr: 0.0001	evaluation reward: 1.6		
episode: 220	score: 2.0	memory length: 41969	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.61		
episode: 221	score: 3.0	memory length: 42194	epsilon: 1.0	steps:
225	lr: 0.0001	evaluation reward: 1.6		
episode: 222	score: 0.0	memory length: 42317	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.58		
episode: 223	score: 1.0	memory length: 42468	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.58		
episode: 224	score: 3.0	memory length: 42699	epsilon: 1.0	steps:
231	lr: 0.0001	evaluation reward: 1.61		
episode: 225	score: 4.0	memory length: 42992	epsilon: 1.0	steps:
293	lr: 0.0001	evaluation reward: 1.65		
episode: 226	score: 3.0	memory length: 43261	epsilon: 1.0	steps:
269	lr: 0.0001	evaluation reward: 1.66		
episode: 227	score: 2.0	memory length: 43460	epsilon: 1.0	steps:
199	lr: 0.0001	evaluation reward: 1.65		
episode: 228	score: 0.0	memory length: 43583	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.64		
episode: 229	score: 1.0	memory length: 43734	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.64		
episode: 230	score: 1.0	memory length: 43885	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.64		
episode: 231	score: 2.0	memory length: 44086	epsilon: 1.0	steps:
201	lr: 0.0001	evaluation reward: 1.66		
episode: 232	score: 2.0	memory length: 44284	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.67		
episode: 233	score: 0.0	memory length: 44407	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.65		
episode: 234	score: 1.0	memory length: 44577	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.65		
episode: 235	score: 4.0	memory length: 44853	epsilon: 1.0	steps:
276	lr: 0.0001	evaluation reward: 1.67		
episode: 236	score: 2.0	memory length: 45056	epsilon: 1.0	steps:
203	lr: 0.0001	evaluation reward: 1.67		
episode: 237	score: 1.0	memory length: 45207	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.66		
episode: 238	score: 1.0	memory length: 45375	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.65		
episode: 239	score: 2.0	memory length: 45593	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.65		
episode: 240	score: 1.0	memory length: 45743	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.65		
episode: 241	score: 2.0	memory length: 45941	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.65		

episode: 242	score: 3.0	memory length: 46190	epsilon: 1.0	steps:
249	lr: 0.0001	evaluation reward: 1.68		
episode: 243	score: 0.0	memory length: 46313	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.67		
episode: 244	score: 0.0	memory length: 46436	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.65		
episode: 245	score: 2.0	memory length: 46634	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.65		
episode: 246	score: 0.0	memory length: 46757	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.64		
episode: 247	score: 2.0	memory length: 46954	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.65		
episode: 248	score: 5.0	memory length: 47289	epsilon: 1.0	steps:
335	lr: 0.0001	evaluation reward: 1.67		
episode: 249	score: 1.0	memory length: 47461	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.67		
episode: 250	score: 2.0	memory length: 47676	epsilon: 1.0	steps:
215	lr: 0.0001	evaluation reward: 1.66		
episode: 251	score: 1.0	memory length: 47846	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.66		
episode: 252	score: 0.0	memory length: 47969	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.65		
episode: 253	score: 2.0	memory length: 48167	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.67		
episode: 254	score: 3.0	memory length: 48412	epsilon: 1.0	steps:
245	lr: 0.0001	evaluation reward: 1.66		
episode: 255	score: 0.0	memory length: 48534	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.66		
episode: 256	score: 0.0	memory length: 48657	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.65		
episode: 257	score: 1.0	memory length: 48807	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.66		
episode: 258	score: 0.0	memory length: 48930	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.66		
episode: 259	score: 1.0	memory length: 49081	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.65		
episode: 260	score: 1.0	memory length: 49249	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.66		
episode: 261	score: 1.0	memory length: 49399	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.65		
episode: 262	score: 2.0	memory length: 49597	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.63		
episode: 263	score: 0.0	memory length: 49720	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.63		
episode: 264	score: 4.0	memory length: 50017	epsilon: 1.0	steps:
297	lr: 0.0001	evaluation reward: 1.65		
episode: 265	score: 4.0	memory length: 50313	epsilon: 1.0	steps:
296	lr: 0.0001	evaluation reward: 1.68		
episode: 266	score: 3.0	memory length: 50543	epsilon: 1.0	steps:
230	lr: 0.0001	evaluation reward: 1.69		
episode: 267	score: 0.0	memory length: 50665	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.67		
episode: 268	score: 0.0	memory length: 50787	epsilon: 1.0	steps:

122	lr: 0.0001	evaluation reward: 1.66		
episode: 269	score: 3.0	memory length: 51034	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.63		
episode: 270	score: 2.0	memory length: 51231	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.65		
episode: 271	score: 1.0	memory length: 51400	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.64		
episode: 272	score: 1.0	memory length: 51571	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.65		
episode: 273	score: 3.0	memory length: 51838	epsilon: 1.0	steps:
267	lr: 0.0001	evaluation reward: 1.64		
episode: 274	score: 2.0	memory length: 52035	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.64		
episode: 275	score: 0.0	memory length: 52158	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.63		
episode: 276	score: 1.0	memory length: 52327	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.64		
episode: 277	score: 0.0	memory length: 52450	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.62		
episode: 278	score: 1.0	memory length: 52601	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.61		
episode: 279	score: 1.0	memory length: 52752	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.62		
episode: 280	score: 4.0	memory length: 53038	epsilon: 1.0	steps:
286	lr: 0.0001	evaluation reward: 1.64		
episode: 281	score: 1.0	memory length: 53188	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.64		
episode: 282	score: 2.0	memory length: 53390	epsilon: 1.0	steps:
202	lr: 0.0001	evaluation reward: 1.66		
episode: 283	score: 3.0	memory length: 53657	epsilon: 1.0	steps:
267	lr: 0.0001	evaluation reward: 1.66		
episode: 284	score: 1.0	memory length: 53808	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.63		
episode: 285	score: 0.0	memory length: 53931	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.63		
episode: 286	score: 1.0	memory length: 54082	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.64		
episode: 287	score: 2.0	memory length: 54300	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.62		
episode: 288	score: 1.0	memory length: 54472	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.62		
episode: 289	score: 3.0	memory length: 54717	epsilon: 1.0	steps:
245	lr: 0.0001	evaluation reward: 1.65		
episode: 290	score: 0.0	memory length: 54840	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.58		
episode: 291	score: 6.0	memory length: 55188	epsilon: 1.0	steps:
348	lr: 0.0001	evaluation reward: 1.63		
episode: 292	score: 3.0	memory length: 55455	epsilon: 1.0	steps:
267	lr: 0.0001	evaluation reward: 1.66		
episode: 293	score: 1.0	memory length: 55606	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.64		
episode: 294	score: 1.0	memory length: 55757	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.64		

episode: 295	score: 1.0	memory length: 55925	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.65		
episode: 296	score: 5.0	memory length: 56231	epsilon: 1.0	steps:
306	lr: 0.0001	evaluation reward: 1.69		
episode: 297	score: 1.0	memory length: 56400	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.68		
episode: 298	score: 1.0	memory length: 56568	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.69		
episode: 299	score: 2.0	memory length: 56765	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.69		
episode: 300	score: 2.0	memory length: 56962	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.7		
episode: 301	score: 2.0	memory length: 57162	epsilon: 1.0	steps:
200	lr: 0.0001	evaluation reward: 1.72		
episode: 302	score: 0.0	memory length: 57285	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.72		
episode: 303	score: 2.0	memory length: 57482	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.72		
episode: 304	score: 1.0	memory length: 57651	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.71		
episode: 305	score: 3.0	memory length: 57877	epsilon: 1.0	steps:
226	lr: 0.0001	evaluation reward: 1.7		
episode: 306	score: 1.0	memory length: 58028	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.68		
episode: 307	score: 0.0	memory length: 58151	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.66		
episode: 308	score: 2.0	memory length: 58332	epsilon: 1.0	steps:
181	lr: 0.0001	evaluation reward: 1.66		
episode: 309	score: 4.0	memory length: 58587	epsilon: 1.0	steps:
255	lr: 0.0001	evaluation reward: 1.67		
episode: 310	score: 3.0	memory length: 58815	epsilon: 1.0	steps:
228	lr: 0.0001	evaluation reward: 1.67		
episode: 311	score: 0.0	memory length: 58938	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.64		
episode: 312	score: 2.0	memory length: 59136	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.64		
episode: 313	score: 4.0	memory length: 59432	epsilon: 1.0	steps:
296	lr: 0.0001	evaluation reward: 1.67		
episode: 314	score: 3.0	memory length: 59702	epsilon: 1.0	steps:
270	lr: 0.0001	evaluation reward: 1.68		
episode: 315	score: 1.0	memory length: 59872	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.68		
episode: 316	score: 2.0	memory length: 60089	epsilon: 1.0	steps:
217	lr: 0.0001	evaluation reward: 1.67		
episode: 317	score: 0.0	memory length: 60212	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.67		
episode: 318	score: 1.0	memory length: 60381	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.67		
episode: 319	score: 1.0	memory length: 60549	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.64		
episode: 320	score: 1.0	memory length: 60717	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.63		
episode: 321	score: 0.0	memory length: 60839	epsilon: 1.0	steps:

122	lr: 0.0001	evaluation reward: 1.6		
episode: 322	score: 3.0	memory length: 61088	epsilon: 1.0	steps:
249	lr: 0.0001	evaluation reward: 1.63		
episode: 323	score: 1.0	memory length: 61257	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.63		
episode: 324	score: 1.0	memory length: 61426	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.61		
episode: 325	score: 6.0	memory length: 61808	epsilon: 1.0	steps:
382	lr: 0.0001	evaluation reward: 1.63		
episode: 326	score: 1.0	memory length: 61959	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.61		
episode: 327	score: 7.0	memory length: 62372	epsilon: 1.0	steps:
413	lr: 0.0001	evaluation reward: 1.66		
episode: 328	score: 0.0	memory length: 62495	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.66		
episode: 329	score: 3.0	memory length: 62759	epsilon: 1.0	steps:
264	lr: 0.0001	evaluation reward: 1.68		
episode: 330	score: 1.0	memory length: 62909	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.68		
episode: 331	score: 0.0	memory length: 63031	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.66		
episode: 332	score: 3.0	memory length: 63259	epsilon: 1.0	steps:
228	lr: 0.0001	evaluation reward: 1.67		
episode: 333	score: 3.0	memory length: 63523	epsilon: 1.0	steps:
264	lr: 0.0001	evaluation reward: 1.7		
episode: 334	score: 2.0	memory length: 63721	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.71		
episode: 335	score: 2.0	memory length: 63919	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.69		
episode: 336	score: 1.0	memory length: 64089	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.68		
episode: 337	score: 2.0	memory length: 64287	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.69		
episode: 338	score: 1.0	memory length: 64458	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.69		
episode: 339	score: 2.0	memory length: 64655	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.69		
episode: 340	score: 1.0	memory length: 64825	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.69		
episode: 341	score: 1.0	memory length: 64996	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.68		
episode: 342	score: 1.0	memory length: 65167	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.66		
episode: 343	score: 3.0	memory length: 65411	epsilon: 1.0	steps:
244	lr: 0.0001	evaluation reward: 1.69		
episode: 344	score: 2.0	memory length: 65609	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.71		
episode: 345	score: 0.0	memory length: 65732	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.69		
episode: 346	score: 0.0	memory length: 65855	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.69		
episode: 347	score: 1.0	memory length: 66006	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.68		

episode: 348	score: 3.0	memory length: 66235	epsilon: 1.0	steps:
229	lr: 0.0001	evaluation reward: 1.66		
episode: 349	score: 1.0	memory length: 66404	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.66		
episode: 350	score: 0.0	memory length: 66526	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.64		
episode: 351	score: 1.0	memory length: 66678	epsilon: 1.0	steps:
152	lr: 0.0001	evaluation reward: 1.64		
episode: 352	score: 3.0	memory length: 66925	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.67		
episode: 353	score: 3.0	memory length: 67150	epsilon: 1.0	steps:
225	lr: 0.0001	evaluation reward: 1.68		
episode: 354	score: 2.0	memory length: 67347	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.67		
episode: 355	score: 2.0	memory length: 67564	epsilon: 1.0	steps:
217	lr: 0.0001	evaluation reward: 1.69		
episode: 356	score: 2.0	memory length: 67744	epsilon: 1.0	steps:
180	lr: 0.0001	evaluation reward: 1.71		
episode: 357	score: 0.0	memory length: 67867	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.7		
episode: 358	score: 1.0	memory length: 68038	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.71		
episode: 359	score: 2.0	memory length: 68235	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.72		
episode: 360	score: 0.0	memory length: 68358	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.71		
episode: 361	score: 0.0	memory length: 68481	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.7		
episode: 362	score: 0.0	memory length: 68604	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.68		
episode: 363	score: 3.0	memory length: 68872	epsilon: 1.0	steps:
268	lr: 0.0001	evaluation reward: 1.71		
episode: 364	score: 1.0	memory length: 69023	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.68		
episode: 365	score: 2.0	memory length: 69221	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.66		
episode: 366	score: 3.0	memory length: 69450	epsilon: 1.0	steps:
229	lr: 0.0001	evaluation reward: 1.66		
episode: 367	score: 1.0	memory length: 69601	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.67		
episode: 368	score: 1.0	memory length: 69771	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.68		
episode: 369	score: 1.0	memory length: 69939	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.66		
episode: 370	score: 1.0	memory length: 70109	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.65		
episode: 371	score: 5.0	memory length: 70415	epsilon: 1.0	steps:
306	lr: 0.0001	evaluation reward: 1.69		
episode: 372	score: 3.0	memory length: 70683	epsilon: 1.0	steps:
268	lr: 0.0001	evaluation reward: 1.71		
episode: 373	score: 0.0	memory length: 70806	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.68		
episode: 374	score: 0.0	memory length: 70928	epsilon: 1.0	steps:

122	lr: 0.0001	evaluation reward: 1.66		
episode: 375	score: 0.0	memory length: 71051	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.66		
episode: 376	score: 1.0	memory length: 71221	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.66		
episode: 377	score: 0.0	memory length: 71344	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.66		
episode: 378	score: 0.0	memory length: 71467	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.65		
episode: 379	score: 0.0	memory length: 71590	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.64		
episode: 380	score: 3.0	memory length: 71817	epsilon: 1.0	steps:
227	lr: 0.0001	evaluation reward: 1.63		
episode: 381	score: 2.0	memory length: 71998	epsilon: 1.0	steps:
181	lr: 0.0001	evaluation reward: 1.64		
episode: 382	score: 0.0	memory length: 72121	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.62		
episode: 383	score: 0.0	memory length: 72243	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.59		
episode: 384	score: 1.0	memory length: 72394	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.59		
episode: 385	score: 2.0	memory length: 72592	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.61		
episode: 386	score: 4.0	memory length: 72889	epsilon: 1.0	steps:
297	lr: 0.0001	evaluation reward: 1.64		
episode: 387	score: 0.0	memory length: 73012	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.62		
episode: 388	score: 1.0	memory length: 73164	epsilon: 1.0	steps:
152	lr: 0.0001	evaluation reward: 1.62		
episode: 389	score: 0.0	memory length: 73287	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.59		
episode: 390	score: 0.0	memory length: 73410	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.59		
episode: 391	score: 4.0	memory length: 73699	epsilon: 1.0	steps:
289	lr: 0.0001	evaluation reward: 1.57		
episode: 392	score: 1.0	memory length: 73869	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.55		
episode: 393	score: 2.0	memory length: 74069	epsilon: 1.0	steps:
200	lr: 0.0001	evaluation reward: 1.56		
episode: 394	score: 1.0	memory length: 74240	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.56		
episode: 395	score: 2.0	memory length: 74458	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.57		
episode: 396	score: 1.0	memory length: 74609	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.53		
episode: 397	score: 2.0	memory length: 74827	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.54		
episode: 398	score: 1.0	memory length: 74998	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.54		
episode: 399	score: 0.0	memory length: 75121	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.52		
episode: 400	score: 1.0	memory length: 75272	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.51		

episode: 401	score: 0.0	memory length: 75395	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.49		
episode: 402	score: 1.0	memory length: 75565	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.5		
episode: 403	score: 3.0	memory length: 75796	epsilon: 1.0	steps:
231	lr: 0.0001	evaluation reward: 1.51		
episode: 404	score: 3.0	memory length: 76044	epsilon: 1.0	steps:
248	lr: 0.0001	evaluation reward: 1.53		
episode: 405	score: 2.0	memory length: 76242	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.52		
episode: 406	score: 1.0	memory length: 76412	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.52		
episode: 407	score: 2.0	memory length: 76614	epsilon: 1.0	steps:
202	lr: 0.0001	evaluation reward: 1.54		
episode: 408	score: 0.0	memory length: 76737	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.52		
episode: 409	score: 1.0	memory length: 76905	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.49		
episode: 410	score: 1.0	memory length: 77076	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.47		
episode: 411	score: 2.0	memory length: 77274	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.49		
episode: 412	score: 0.0	memory length: 77396	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.47		
episode: 413	score: 0.0	memory length: 77519	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.43		
episode: 414	score: 5.0	memory length: 77861	epsilon: 1.0	steps:
342	lr: 0.0001	evaluation reward: 1.45		
episode: 415	score: 0.0	memory length: 77983	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.44		
episode: 416	score: 0.0	memory length: 78105	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.42		
episode: 417	score: 0.0	memory length: 78227	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.42		
episode: 418	score: 1.0	memory length: 78378	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.42		
episode: 419	score: 0.0	memory length: 78501	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.41		
episode: 420	score: 0.0	memory length: 78624	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.4		
episode: 421	score: 1.0	memory length: 78793	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.41		
episode: 422	score: 0.0	memory length: 78916	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.38		
episode: 423	score: 0.0	memory length: 79039	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.37		
episode: 424	score: 3.0	memory length: 79289	epsilon: 1.0	steps:
250	lr: 0.0001	evaluation reward: 1.39		
episode: 425	score: 3.0	memory length: 79533	epsilon: 1.0	steps:
244	lr: 0.0001	evaluation reward: 1.36		
episode: 426	score: 0.0	memory length: 79656	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.35		
episode: 427	score: 0.0	memory length: 79779	epsilon: 1.0	steps:

123	lr: 0.0001	evaluation reward: 1.28		
episode: 428	score: 0.0	memory length: 79902	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.28		
episode: 429	score: 3.0	memory length: 80152	epsilon: 1.0	steps:
250	lr: 0.0001	evaluation reward: 1.28		
episode: 430	score: 1.0	memory length: 80320	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.28		
episode: 431	score: 0.0	memory length: 80443	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.28		
episode: 432	score: 3.0	memory length: 80671	epsilon: 1.0	steps:
228	lr: 0.0001	evaluation reward: 1.28		
episode: 433	score: 1.0	memory length: 80823	epsilon: 1.0	steps:
152	lr: 0.0001	evaluation reward: 1.26		
episode: 434	score: 0.0	memory length: 80946	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.24		
episode: 435	score: 0.0	memory length: 81069	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.22		
episode: 436	score: 2.0	memory length: 81287	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.23		
episode: 437	score: 4.0	memory length: 81584	epsilon: 1.0	steps:
297	lr: 0.0001	evaluation reward: 1.25		
episode: 438	score: 2.0	memory length: 81786	epsilon: 1.0	steps:
202	lr: 0.0001	evaluation reward: 1.26		
episode: 439	score: 1.0	memory length: 81955	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.25		
episode: 440	score: 2.0	memory length: 82154	epsilon: 1.0	steps:
199	lr: 0.0001	evaluation reward: 1.26		
episode: 441	score: 0.0	memory length: 82276	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.25		
episode: 442	score: 2.0	memory length: 82494	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.26		
episode: 443	score: 1.0	memory length: 82663	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.24		
episode: 444	score: 0.0	memory length: 82786	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.22		
episode: 445	score: 2.0	memory length: 83004	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.24		
episode: 446	score: 4.0	memory length: 83259	epsilon: 1.0	steps:
255	lr: 0.0001	evaluation reward: 1.28		
episode: 447	score: 2.0	memory length: 83441	epsilon: 1.0	steps:
182	lr: 0.0001	evaluation reward: 1.29		
episode: 448	score: 1.0	memory length: 83592	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.27		
episode: 449	score: 2.0	memory length: 83790	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.28		
episode: 450	score: 0.0	memory length: 83912	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.28		
episode: 451	score: 0.0	memory length: 84035	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.27		
episode: 452	score: 2.0	memory length: 84233	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.26		
episode: 453	score: 0.0	memory length: 84356	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.23		

episode: 454	score: 1.0	memory length: 84525	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.22		
episode: 455	score: 1.0	memory length: 84694	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.21		
episode: 456	score: 2.0	memory length: 84912	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.21		
episode: 457	score: 1.0	memory length: 85081	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.22		
episode: 458	score: 0.0	memory length: 85203	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.21		
episode: 459	score: 2.0	memory length: 85390	epsilon: 1.0	steps:
187	lr: 0.0001	evaluation reward: 1.21		
episode: 460	score: 0.0	memory length: 85513	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.21		
episode: 461	score: 2.0	memory length: 85710	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.23		
episode: 462	score: 1.0	memory length: 85860	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.24		
episode: 463	score: 3.0	memory length: 86088	epsilon: 1.0	steps:
228	lr: 0.0001	evaluation reward: 1.24		
episode: 464	score: 0.0	memory length: 86211	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.23		
episode: 465	score: 4.0	memory length: 86525	epsilon: 1.0	steps:
314	lr: 0.0001	evaluation reward: 1.25		
episode: 466	score: 2.0	memory length: 86722	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.24		
episode: 467	score: 1.0	memory length: 86891	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.24		
episode: 468	score: 0.0	memory length: 87014	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.23		
episode: 469	score: 1.0	memory length: 87185	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.23		
episode: 470	score: 1.0	memory length: 87353	epsilon: 1.0	steps:
168	lr: 0.0001	evaluation reward: 1.23		
episode: 471	score: 5.0	memory length: 87700	epsilon: 1.0	steps:
347	lr: 0.0001	evaluation reward: 1.23		
episode: 472	score: 3.0	memory length: 87947	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.23		
episode: 473	score: 2.0	memory length: 88145	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.25		
episode: 474	score: 0.0	memory length: 88267	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.25		
episode: 475	score: 1.0	memory length: 88437	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.26		
episode: 476	score: 3.0	memory length: 88702	epsilon: 1.0	steps:
265	lr: 0.0001	evaluation reward: 1.28		
episode: 477	score: 0.0	memory length: 88824	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.28		
episode: 478	score: 2.0	memory length: 89044	epsilon: 1.0	steps:
220	lr: 0.0001	evaluation reward: 1.3		
episode: 479	score: 2.0	memory length: 89242	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.32		
episode: 480	score: 3.0	memory length: 89452	epsilon: 1.0	steps:

210	lr: 0.0001	evaluation reward: 1.32		
episode: 481	score: 2.0	memory length: 89649	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.32		
episode: 482	score: 4.0	memory length: 89925	epsilon: 1.0	steps:
276	lr: 0.0001	evaluation reward: 1.36		
episode: 483	score: 3.0	memory length: 90190	epsilon: 1.0	steps:
265	lr: 0.0001	evaluation reward: 1.39		
episode: 484	score: 2.0	memory length: 90406	epsilon: 1.0	steps:
216	lr: 0.0001	evaluation reward: 1.4		
episode: 485	score: 2.0	memory length: 90586	epsilon: 1.0	steps:
180	lr: 0.0001	evaluation reward: 1.4		
episode: 486	score: 4.0	memory length: 90852	epsilon: 1.0	steps:
266	lr: 0.0001	evaluation reward: 1.4		
episode: 487	score: 1.0	memory length: 91002	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.41		
episode: 488	score: 0.0	memory length: 91124	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.4		
episode: 489	score: 3.0	memory length: 91368	epsilon: 1.0	steps:
244	lr: 0.0001	evaluation reward: 1.43		
episode: 490	score: 1.0	memory length: 91538	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.44		
episode: 491	score: 4.0	memory length: 91816	epsilon: 1.0	steps:
278	lr: 0.0001	evaluation reward: 1.44		
episode: 492	score: 1.0	memory length: 91985	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.44		
episode: 493	score: 5.0	memory length: 92323	epsilon: 1.0	steps:
338	lr: 0.0001	evaluation reward: 1.47		
episode: 494	score: 3.0	memory length: 92552	epsilon: 1.0	steps:
229	lr: 0.0001	evaluation reward: 1.49		
episode: 495	score: 4.0	memory length: 92845	epsilon: 1.0	steps:
293	lr: 0.0001	evaluation reward: 1.51		
episode: 496	score: 1.0	memory length: 93017	epsilon: 1.0	steps:
172	lr: 0.0001	evaluation reward: 1.51		
episode: 497	score: 0.0	memory length: 93140	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.49		
episode: 498	score: 0.0	memory length: 93263	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.48		
episode: 499	score: 0.0	memory length: 93386	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.48		
episode: 500	score: 0.0	memory length: 93509	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.47		
episode: 501	score: 2.0	memory length: 93706	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.49		
episode: 502	score: 3.0	memory length: 93953	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.51		
episode: 503	score: 1.0	memory length: 94123	epsilon: 1.0	steps:
170	lr: 0.0001	evaluation reward: 1.49		
episode: 504	score: 4.0	memory length: 94405	epsilon: 1.0	steps:
282	lr: 0.0001	evaluation reward: 1.5		
episode: 505	score: 0.0	memory length: 94528	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.48		
episode: 506	score: 0.0	memory length: 94651	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.47		

episode: 507	score: 1.0	memory length: 94820	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.46		
episode: 508	score: 0.0	memory length: 94943	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.46		
episode: 509	score: 3.0	memory length: 95188	epsilon: 1.0	steps:
245	lr: 0.0001	evaluation reward: 1.48		
episode: 510	score: 3.0	memory length: 95455	epsilon: 1.0	steps:
267	lr: 0.0001	evaluation reward: 1.5		
episode: 511	score: 1.0	memory length: 95626	epsilon: 1.0	steps:
171	lr: 0.0001	evaluation reward: 1.49		
episode: 512	score: 1.0	memory length: 95777	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.5		
episode: 513	score: 1.0	memory length: 95928	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.51		
episode: 514	score: 2.0	memory length: 96146	epsilon: 1.0	steps:
218	lr: 0.0001	evaluation reward: 1.48		
episode: 515	score: 1.0	memory length: 96297	epsilon: 1.0	steps:
151	lr: 0.0001	evaluation reward: 1.49		
episode: 516	score: 3.0	memory length: 96563	epsilon: 1.0	steps:
266	lr: 0.0001	evaluation reward: 1.52		
episode: 517	score: 1.0	memory length: 96732	epsilon: 1.0	steps:
169	lr: 0.0001	evaluation reward: 1.53		
episode: 518	score: 3.0	memory length: 96978	epsilon: 1.0	steps:
246	lr: 0.0001	evaluation reward: 1.55		
episode: 519	score: 0.0	memory length: 97101	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.55		
episode: 520	score: 0.0	memory length: 97224	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.55		
episode: 521	score: 2.0	memory length: 97421	epsilon: 1.0	steps:
197	lr: 0.0001	evaluation reward: 1.56		
episode: 522	score: 1.0	memory length: 97571	epsilon: 1.0	steps:
150	lr: 0.0001	evaluation reward: 1.57		
episode: 523	score: 0.0	memory length: 97694	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.57		
episode: 524	score: 2.0	memory length: 97894	epsilon: 1.0	steps:
200	lr: 0.0001	evaluation reward: 1.56		
episode: 525	score: 2.0	memory length: 98074	epsilon: 1.0	steps:
180	lr: 0.0001	evaluation reward: 1.55		
episode: 526	score: 3.0	memory length: 98339	epsilon: 1.0	steps:
265	lr: 0.0001	evaluation reward: 1.58		
episode: 527	score: 0.0	memory length: 98462	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.58		
episode: 528	score: 0.0	memory length: 98584	epsilon: 1.0	steps:
122	lr: 0.0001	evaluation reward: 1.58		
episode: 529	score: 2.0	memory length: 98782	epsilon: 1.0	steps:
198	lr: 0.0001	evaluation reward: 1.57		
episode: 530	score: 0.0	memory length: 98905	epsilon: 1.0	steps:
123	lr: 0.0001	evaluation reward: 1.56		
episode: 531	score: 3.0	memory length: 99152	epsilon: 1.0	steps:
247	lr: 0.0001	evaluation reward: 1.59		
episode: 532	score: 2.0	memory length: 99372	epsilon: 1.0	steps:
220	lr: 0.0001	evaluation reward: 1.58		
episode: 533	score: 0.0	memory length: 99495	epsilon: 1.0	steps:

```

123    lr: 0.0001    evaluation reward: 1.57
episode: 534    score: 0.0    memory length: 99618    epsilon: 1.0    steps:
123    lr: 0.0001    evaluation reward: 1.57
episode: 535    score: 1.0    memory length: 99787    epsilon: 1.0    steps:
169    lr: 0.0001    evaluation reward: 1.58
episode: 536    score: 1.0    memory length: 99957    epsilon: 1.0    steps:
170    lr: 0.0001    evaluation reward: 1.57

```

/home/wxzhang/assignment5/memory.py:30: FutureWarning: The input object of type 'Tensor' is an array-like implementing one of the corresponding protocols ('__array__', '__array_interface__' or '__array_struct__'); but not a sequence (or 0-D). In the future, this object will be coerced as if it was first converted using 'np.array(obj)'. To retain the old behaviour, you have to either modify the type 'Tensor', or assign to an empty array created with 'np.empty(correct_shape, dtype=object)'.

```
sample = np.array(sample, dtype=object)
```

/home/wxzhang/assignment5/agent_double.py:78: FutureWarning: The input object of type 'Tensor' is an array-like implementing one of the corresponding protocols ('__array__', '__array_interface__' or '__array_struct__'); but not a sequence (or 0-D). In the future, this object will be coerced as if it was first converted using 'np.array(obj)'. To retain the old behaviour, you have to either modify the type 'Tensor', or assign to an empty array created with 'np.empty(correct_shape, dtype=object)'.

```
mini_batch = np.array(mini_batch, dtype=object).transpose()
```

```

episode: 537    score: 1.0    memory length: 100108    epsilon: 0.99978418000
00047    steps: 151    lr: 0.0001    evaluation reward: 1.54
episode: 538    score: 1.0    memory length: 100278    epsilon: 0.99944758000
0012    steps: 170    lr: 0.0001    evaluation reward: 1.53
episode: 539    score: 0.0    memory length: 100401    epsilon: 0.99920404000
00173    steps: 123    lr: 0.0001    evaluation reward: 1.52
episode: 540    score: 3.0    memory length: 100631    epsilon: 0.99874864000
00272    steps: 230    lr: 0.0001    evaluation reward: 1.53
episode: 541    score: 1.0    memory length: 100802    epsilon: 0.99841006000
00345    steps: 171    lr: 0.0001    evaluation reward: 1.54
episode: 542    score: 1.0    memory length: 100952    epsilon: 0.99811306000
0041    steps: 150    lr: 0.0001    evaluation reward: 1.53
episode: 543    score: 1.0    memory length: 101103    epsilon: 0.99781408000
00475    steps: 151    lr: 0.0001    evaluation reward: 1.53
episode: 544    score: 7.0    memory length: 101513    epsilon: 0.99700228000
00651    steps: 410    lr: 0.0001    evaluation reward: 1.6
episode: 545    score: 3.0    memory length: 101763    epsilon: 0.99650728000
00758    steps: 250    lr: 0.0001    evaluation reward: 1.61
episode: 546    score: 1.0    memory length: 101932    epsilon: 0.99617266000
00831    steps: 169    lr: 0.0001    evaluation reward: 1.58
episode: 547    score: 1.0    memory length: 102084    epsilon: 0.99587170000
00896    steps: 152    lr: 0.0001    evaluation reward: 1.57
episode: 548    score: 2.0    memory length: 102300    epsilon: 0.99544402000
00989    steps: 216    lr: 0.0001    evaluation reward: 1.58
episode: 549    score: 5.0    memory length: 102627    epsilon: 0.99479656000
0113    steps: 327    lr: 0.0001    evaluation reward: 1.61
episode: 550    score: 2.0    memory length: 102844    epsilon: 0.99436690000
01223    steps: 217    lr: 0.0001    evaluation reward: 1.63
episode: 551    score: 0.0    memory length: 102966    epsilon: 0.99412534000
01275    steps: 122    lr: 0.0001    evaluation reward: 1.63

```

```
episode: 552  score: 1.0  memory length: 103135  epsilon: 0.99379072000
01348  steps: 169  lr: 0.0001  evaluation reward: 1.62
episode: 553  score: 1.0  memory length: 103307  epsilon: 0.99345016000
01422  steps: 172  lr: 0.0001  evaluation reward: 1.63
episode: 554  score: 0.0  memory length: 103430  epsilon: 0.99320662000
01475  steps: 123  lr: 0.0001  evaluation reward: 1.62
episode: 555  score: 2.0  memory length: 103648  epsilon: 0.99277498000
01568  steps: 218  lr: 0.0001  evaluation reward: 1.63
episode: 556  score: 2.0  memory length: 103828  epsilon: 0.99241858000
01646  steps: 180  lr: 0.0001  evaluation reward: 1.63
episode: 557  score: 0.0  memory length: 103951  epsilon: 0.99217504000
01699  steps: 123  lr: 0.0001  evaluation reward: 1.62
episode: 558  score: 3.0  memory length: 104218  epsilon: 0.99164638000
01813  steps: 267  lr: 0.0001  evaluation reward: 1.65
episode: 559  score: 3.0  memory length: 104465  epsilon: 0.99115732000
0192  steps: 247  lr: 0.0001  evaluation reward: 1.66
episode: 560  score: 2.0  memory length: 104663  epsilon: 0.99076528000
02005  steps: 198  lr: 0.0001  evaluation reward: 1.68
episode: 561  score: 1.0  memory length: 104814  epsilon: 0.99046630000
0207  steps: 151  lr: 0.0001  evaluation reward: 1.67
episode: 562  score: 4.0  memory length: 105089  epsilon: 0.98992180000
02188  steps: 275  lr: 0.0001  evaluation reward: 1.7
episode: 563  score: 3.0  memory length: 105320  epsilon: 0.98946442000
02287  steps: 231  lr: 0.0001  evaluation reward: 1.7
episode: 564  score: 1.0  memory length: 105471  epsilon: 0.98916544000
02352  steps: 151  lr: 0.0001  evaluation reward: 1.71
episode: 565  score: 0.0  memory length: 105594  epsilon: 0.98892190000
02405  steps: 123  lr: 0.0001  evaluation reward: 1.67
episode: 566  score: 0.0  memory length: 105717  epsilon: 0.98867836000
02458  steps: 123  lr: 0.0001  evaluation reward: 1.65
episode: 567  score: 4.0  memory length: 106027  epsilon: 0.98806456000
02591  steps: 310  lr: 0.0001  evaluation reward: 1.68
episode: 568  score: 2.0  memory length: 106248  epsilon: 0.98762698000
02686  steps: 221  lr: 0.0001  evaluation reward: 1.7
episode: 569  score: 0.0  memory length: 106371  epsilon: 0.98738344000
02739  steps: 123  lr: 0.0001  evaluation reward: 1.69
episode: 570  score: 5.0  memory length: 106667  epsilon: 0.98679736000
02866  steps: 296  lr: 0.0001  evaluation reward: 1.73
episode: 571  score: 3.0  memory length: 106932  epsilon: 0.98627266000
0298  steps: 265  lr: 0.0001  evaluation reward: 1.71
episode: 572  score: 5.0  memory length: 107297  epsilon: 0.98554996000
03137  steps: 365  lr: 0.0001  evaluation reward: 1.73
episode: 573  score: 0.0  memory length: 107420  epsilon: 0.98530642000
0319  steps: 123  lr: 0.0001  evaluation reward: 1.71
episode: 574  score: 2.0  memory length: 107636  epsilon: 0.98487874000
03283  steps: 216  lr: 0.0001  evaluation reward: 1.73
episode: 575  score: 3.0  memory length: 107884  epsilon: 0.98438770000
03389  steps: 248  lr: 0.0001  evaluation reward: 1.75
episode: 576  score: 3.0  memory length: 108131  epsilon: 0.98389864000
03495  steps: 247  lr: 0.0001  evaluation reward: 1.75
episode: 577  score: 0.0  memory length: 108254  epsilon: 0.98365510000
03548  steps: 123  lr: 0.0001  evaluation reward: 1.75
episode: 578  score: 2.0  memory length: 108454  epsilon: 0.98325910000
```

```
03634  steps: 200  lr: 0.0001  evaluation reward: 1.75
episode: 579  score: 1.0  memory length: 108605  epsilon: 0.98296012000
03699  steps: 151  lr: 0.0001  evaluation reward: 1.74
episode: 580  score: 2.0  memory length: 108786  epsilon: 0.98260174000
03777  steps: 181  lr: 0.0001  evaluation reward: 1.73
episode: 581  score: 3.0  memory length: 109002  epsilon: 0.98217406000
0387  steps: 216  lr: 0.0001  evaluation reward: 1.74
episode: 582  score: 1.0  memory length: 109154  epsilon: 0.98187310000
03935  steps: 152  lr: 0.0001  evaluation reward: 1.71
episode: 583  score: 1.0  memory length: 109304  epsilon: 0.98157610000
04  steps: 150  lr: 0.0001  evaluation reward: 1.69
episode: 584  score: 3.0  memory length: 109519  epsilon: 0.98115040000
04092  steps: 215  lr: 0.0001  evaluation reward: 1.7
episode: 585  score: 2.0  memory length: 109717  epsilon: 0.98075836000
04177  steps: 198  lr: 0.0001  evaluation reward: 1.7
episode: 586  score: 1.0  memory length: 109886  epsilon: 0.98042374000
0425  steps: 169  lr: 0.0001  evaluation reward: 1.67
episode: 587  score: 1.0  memory length: 110054  epsilon: 0.98009110000
04322  steps: 168  lr: 0.0001  evaluation reward: 1.67
episode: 588  score: 2.0  memory length: 110252  epsilon: 0.97969906000
04407  steps: 198  lr: 0.0001  evaluation reward: 1.69
episode: 589  score: 1.0  memory length: 110422  epsilon: 0.97936246000
0448  steps: 170  lr: 0.0001  evaluation reward: 1.67
episode: 590  score: 1.0  memory length: 110591  epsilon: 0.97902784000
04553  steps: 169  lr: 0.0001  evaluation reward: 1.67
episode: 591  score: 0.0  memory length: 110713  epsilon: 0.97878628000
04605  steps: 122  lr: 0.0001  evaluation reward: 1.63
episode: 592  score: 2.0  memory length: 110931  epsilon: 0.97835464000
04699  steps: 218  lr: 0.0001  evaluation reward: 1.64
episode: 593  score: 3.0  memory length: 111148  epsilon: 0.97792498000
04792  steps: 217  lr: 0.0001  evaluation reward: 1.62
episode: 594  score: 1.0  memory length: 111299  epsilon: 0.97762600000
04857  steps: 151  lr: 0.0001  evaluation reward: 1.6
episode: 595  score: 0.0  memory length: 111421  epsilon: 0.97738444000
0491  steps: 122  lr: 0.0001  evaluation reward: 1.56
episode: 596  score: 0.0  memory length: 111543  epsilon: 0.97714288000
04962  steps: 122  lr: 0.0001  evaluation reward: 1.55
episode: 597  score: 0.0  memory length: 111665  epsilon: 0.97690132000
05014  steps: 122  lr: 0.0001  evaluation reward: 1.55
episode: 598  score: 4.0  memory length: 111927  epsilon: 0.97638256000
05127  steps: 262  lr: 0.0001  evaluation reward: 1.59
episode: 599  score: 4.0  memory length: 112243  epsilon: 0.97575688000
05263  steps: 316  lr: 0.0001  evaluation reward: 1.63
episode: 600  score: 1.0  memory length: 112394  epsilon: 0.97545790000
05328  steps: 151  lr: 0.0001  evaluation reward: 1.64
episode: 601  score: 0.0  memory length: 112516  epsilon: 0.97521634000
0538  steps: 122  lr: 0.0001  evaluation reward: 1.62
episode: 602  score: 2.0  memory length: 112714  epsilon: 0.97482430000
05465  steps: 198  lr: 0.0001  evaluation reward: 1.61
episode: 603  score: 0.0  memory length: 112837  epsilon: 0.97458076000
05518  steps: 123  lr: 0.0001  evaluation reward: 1.6
episode: 604  score: 1.0  memory length: 112989  epsilon: 0.97427980000
05584  steps: 152  lr: 0.0001  evaluation reward: 1.57
```

```
episode: 605   score: 3.0   memory length: 113256   epsilon: 0.97375114000
05698   steps: 267   lr: 0.0001   evaluation reward: 1.6
episode: 606   score: 1.0   memory length: 113428   epsilon: 0.97341058000
05772   steps: 172   lr: 0.0001   evaluation reward: 1.61
episode: 607   score: 1.0   memory length: 113597   epsilon: 0.97307596000
05845   steps: 169   lr: 0.0001   evaluation reward: 1.61
episode: 608   score: 3.0   memory length: 113842   epsilon: 0.97259086000
0595   steps: 245   lr: 0.0001   evaluation reward: 1.64
episode: 609   score: 2.0   memory length: 114022   epsilon: 0.97223446000
06028   steps: 180   lr: 0.0001   evaluation reward: 1.63
episode: 610   score: 3.0   memory length: 114270   epsilon: 0.97174342000
06134   steps: 248   lr: 0.0001   evaluation reward: 1.63
episode: 611   score: 1.0   memory length: 114440   epsilon: 0.97140682000
06207   steps: 170   lr: 0.0001   evaluation reward: 1.63
episode: 612   score: 1.0   memory length: 114591   epsilon: 0.97110784000
06272   steps: 151   lr: 0.0001   evaluation reward: 1.63
episode: 613   score: 2.0   memory length: 114807   epsilon: 0.97068016000
06365   steps: 216   lr: 0.0001   evaluation reward: 1.64
episode: 614   score: 2.0   memory length: 115005   epsilon: 0.97028812000
0645   steps: 198   lr: 0.0001   evaluation reward: 1.64
episode: 615   score: 0.0   memory length: 115127   epsilon: 0.97004656000
06503   steps: 122   lr: 0.0001   evaluation reward: 1.63
episode: 616   score: 5.0   memory length: 115452   epsilon: 0.96940306000
06642   steps: 325   lr: 0.0001   evaluation reward: 1.65
episode: 617   score: 0.0   memory length: 115575   epsilon: 0.96915952000
06695   steps: 123   lr: 0.0001   evaluation reward: 1.64
episode: 618   score: 2.0   memory length: 115773   epsilon: 0.96876748000
0678   steps: 198   lr: 0.0001   evaluation reward: 1.63
episode: 619   score: 0.0   memory length: 115895   epsilon: 0.96852592000
06833   steps: 122   lr: 0.0001   evaluation reward: 1.63
episode: 620   score: 2.0   memory length: 116114   epsilon: 0.96809230000
06927   steps: 219   lr: 0.0001   evaluation reward: 1.65
episode: 621   score: 1.0   memory length: 116283   epsilon: 0.96775768000
07   steps: 169   lr: 0.0001   evaluation reward: 1.64
episode: 622   score: 0.0   memory length: 116406   epsilon: 0.96751414000
07052   steps: 123   lr: 0.0001   evaluation reward: 1.63
episode: 623   score: 1.0   memory length: 116576   epsilon: 0.96717754000
07125   steps: 170   lr: 0.0001   evaluation reward: 1.64
episode: 624   score: 0.0   memory length: 116698   epsilon: 0.96693598000
07178   steps: 122   lr: 0.0001   evaluation reward: 1.62
episode: 625   score: 1.0   memory length: 116869   epsilon: 0.96659740000
07251   steps: 171   lr: 0.0001   evaluation reward: 1.61
episode: 626   score: 1.0   memory length: 117038   epsilon: 0.96626278000
07324   steps: 169   lr: 0.0001   evaluation reward: 1.59
episode: 627   score: 0.0   memory length: 117160   epsilon: 0.96602122000
07376   steps: 122   lr: 0.0001   evaluation reward: 1.59
episode: 628   score: 0.0   memory length: 117282   epsilon: 0.96577966000
07429   steps: 122   lr: 0.0001   evaluation reward: 1.59
episode: 629   score: 0.0   memory length: 117404   epsilon: 0.96553810000
07481   steps: 122   lr: 0.0001   evaluation reward: 1.57
episode: 630   score: 2.0   memory length: 117602   epsilon: 0.96514606000
07566   steps: 198   lr: 0.0001   evaluation reward: 1.59
episode: 631   score: 0.0   memory length: 117724   epsilon: 0.96490450000
```

```
07619  steps: 122  lr: 0.0001  evaluation reward: 1.56
episode: 632  score: 0.0  memory length: 117847  epsilon: 0.96466096000
07672  steps: 123  lr: 0.0001  evaluation reward: 1.54
episode: 633  score: 1.0  memory length: 118017  epsilon: 0.96432436000
07745  steps: 170  lr: 0.0001  evaluation reward: 1.55
episode: 634  score: 2.0  memory length: 118198  epsilon: 0.96396598000
07823  steps: 181  lr: 0.0001  evaluation reward: 1.57
episode: 635  score: 0.0  memory length: 118321  epsilon: 0.96372244000
07876  steps: 123  lr: 0.0001  evaluation reward: 1.56
episode: 636  score: 0.0  memory length: 118444  epsilon: 0.96347890000
07928  steps: 123  lr: 0.0001  evaluation reward: 1.55
episode: 637  score: 0.0  memory length: 118566  epsilon: 0.96323734000
07981  steps: 122  lr: 0.0001  evaluation reward: 1.54
episode: 638  score: 2.0  memory length: 118786  epsilon: 0.96280174000
08075  steps: 220  lr: 0.0001  evaluation reward: 1.55
episode: 639  score: 7.0  memory length: 119086  epsilon: 0.96220774000
08204  steps: 300  lr: 0.0001  evaluation reward: 1.62
episode: 640  score: 0.0  memory length: 119208  epsilon: 0.96196618000
08257  steps: 122  lr: 0.0001  evaluation reward: 1.59
episode: 641  score: 4.0  memory length: 119522  epsilon: 0.96134446000
08392  steps: 314  lr: 0.0001  evaluation reward: 1.62
episode: 642  score: 0.0  memory length: 119645  epsilon: 0.96110092000
08445  steps: 123  lr: 0.0001  evaluation reward: 1.61
episode: 643  score: 0.0  memory length: 119767  epsilon: 0.96085936000
08497  steps: 122  lr: 0.0001  evaluation reward: 1.6
episode: 644  score: 0.0  memory length: 119890  epsilon: 0.96061582000
0855  steps: 123  lr: 0.0001  evaluation reward: 1.53
episode: 645  score: 2.0  memory length: 120088  epsilon: 0.96022378000
08635  steps: 198  lr: 0.0001  evaluation reward: 1.52
episode: 646  score: 2.0  memory length: 120285  epsilon: 0.95983372000
0872  steps: 197  lr: 0.0001  evaluation reward: 1.53
episode: 647  score: 2.0  memory length: 120504  epsilon: 0.95940010000
08814  steps: 219  lr: 0.0001  evaluation reward: 1.54
episode: 648  score: 0.0  memory length: 120627  epsilon: 0.95915656000
08867  steps: 123  lr: 0.0001  evaluation reward: 1.52
episode: 649  score: 2.0  memory length: 120845  epsilon: 0.95872492000
0896  steps: 218  lr: 0.0001  evaluation reward: 1.49
episode: 650  score: 0.0  memory length: 120968  epsilon: 0.95848138000
09013  steps: 123  lr: 0.0001  evaluation reward: 1.47
episode: 651  score: 0.0  memory length: 121090  epsilon: 0.95823982000
09066  steps: 122  lr: 0.0001  evaluation reward: 1.47
episode: 652  score: 0.0  memory length: 121213  epsilon: 0.95799628000
09119  steps: 123  lr: 0.0001  evaluation reward: 1.46
episode: 653  score: 0.0  memory length: 121336  epsilon: 0.95775274000
09171  steps: 123  lr: 0.0001  evaluation reward: 1.45
episode: 654  score: 0.0  memory length: 121459  epsilon: 0.95750920000
09224  steps: 123  lr: 0.0001  evaluation reward: 1.45
episode: 655  score: 0.0  memory length: 121582  epsilon: 0.95726566000
09277  steps: 123  lr: 0.0001  evaluation reward: 1.43
episode: 656  score: 3.0  memory length: 121810  epsilon: 0.95681422000
09375  steps: 228  lr: 0.0001  evaluation reward: 1.44
episode: 657  score: 2.0  memory length: 122027  epsilon: 0.95638456000
09468  steps: 217  lr: 0.0001  evaluation reward: 1.46
```

```
episode: 658  score: 1.0  memory length: 122196  epsilon: 0.95604994000
09541  steps: 169  lr: 0.0001  evaluation reward: 1.44
episode: 659  score: 2.0  memory length: 122397  epsilon: 0.95565196000
09628  steps: 201  lr: 0.0001  evaluation reward: 1.43
episode: 660  score: 0.0  memory length: 122519  epsilon: 0.95541040000
0968  steps: 122  lr: 0.0001  evaluation reward: 1.41
episode: 661  score: 2.0  memory length: 122737  epsilon: 0.95497876000
09774  steps: 218  lr: 0.0001  evaluation reward: 1.42
episode: 662  score: 3.0  memory length: 122963  epsilon: 0.95453128000
09871  steps: 226  lr: 0.0001  evaluation reward: 1.41
episode: 663  score: 2.0  memory length: 123182  epsilon: 0.95409766000
09965  steps: 219  lr: 0.0001  evaluation reward: 1.4
episode: 664  score: 1.0  memory length: 123350  epsilon: 0.95376502000
10037  steps: 168  lr: 0.0001  evaluation reward: 1.4
episode: 665  score: 0.0  memory length: 123472  epsilon: 0.95352346000
1009  steps: 122  lr: 0.0001  evaluation reward: 1.4
episode: 666  score: 0.0  memory length: 123594  epsilon: 0.95328190000
10142  steps: 122  lr: 0.0001  evaluation reward: 1.4
episode: 667  score: 3.0  memory length: 123837  epsilon: 0.95280076000
10246  steps: 243  lr: 0.0001  evaluation reward: 1.39
episode: 668  score: 2.0  memory length: 124056  epsilon: 0.95236714000
10341  steps: 219  lr: 0.0001  evaluation reward: 1.39
episode: 669  score: 4.0  memory length: 124349  epsilon: 0.95178700000
10467  steps: 293  lr: 0.0001  evaluation reward: 1.43
episode: 670  score: 0.0  memory length: 124472  epsilon: 0.95154346000
10519  steps: 123  lr: 0.0001  evaluation reward: 1.38
episode: 671  score: 3.0  memory length: 124718  epsilon: 0.95105638000
10625  steps: 246  lr: 0.0001  evaluation reward: 1.38
episode: 672  score: 2.0  memory length: 124938  epsilon: 0.95062078000
1072  steps: 220  lr: 0.0001  evaluation reward: 1.35
episode: 673  score: 2.0  memory length: 125137  epsilon: 0.95022676000
10805  steps: 199  lr: 0.0001  evaluation reward: 1.37
episode: 674  score: 2.0  memory length: 125335  epsilon: 0.94983472000
1089  steps: 198  lr: 0.0001  evaluation reward: 1.37
episode: 675  score: 4.0  memory length: 125630  epsilon: 0.94925062000
11017  steps: 295  lr: 0.0001  evaluation reward: 1.38
episode: 676  score: 3.0  memory length: 125877  epsilon: 0.94876156000
11123  steps: 247  lr: 0.0001  evaluation reward: 1.38
episode: 677  score: 0.0  memory length: 125999  epsilon: 0.94852000000
11176  steps: 122  lr: 0.0001  evaluation reward: 1.38
episode: 678  score: 1.0  memory length: 126150  epsilon: 0.94822102000
11241  steps: 151  lr: 0.0001  evaluation reward: 1.37
episode: 679  score: 2.0  memory length: 126351  epsilon: 0.94782304000
11327  steps: 201  lr: 0.0001  evaluation reward: 1.38
episode: 680  score: 0.0  memory length: 126473  epsilon: 0.94758148000
1138  steps: 122  lr: 0.0001  evaluation reward: 1.36
episode: 681  score: 5.0  memory length: 126782  epsilon: 0.94696966000
11512  steps: 309  lr: 0.0001  evaluation reward: 1.38
episode: 682  score: 3.0  memory length: 127029  epsilon: 0.94648060000
11619  steps: 247  lr: 0.0001  evaluation reward: 1.4
episode: 683  score: 3.0  memory length: 127256  epsilon: 0.94603114000
11716  steps: 227  lr: 0.0001  evaluation reward: 1.42
episode: 684  score: 1.0  memory length: 127407  epsilon: 0.94573216000
```

```
11781  steps: 151  lr: 0.0001  evaluation reward: 1.4
episode: 685  score: 3.0  memory length: 127675  epsilon: 0.94520152000
11896  steps: 268  lr: 0.0001  evaluation reward: 1.41
episode: 686  score: 1.0  memory length: 127844  epsilon: 0.94486690000
11969  steps: 169  lr: 0.0001  evaluation reward: 1.41
episode: 687  score: 0.0  memory length: 127967  epsilon: 0.94462336000
12022  steps: 123  lr: 0.0001  evaluation reward: 1.4
episode: 688  score: 2.0  memory length: 128167  epsilon: 0.94422736000
12108  steps: 200  lr: 0.0001  evaluation reward: 1.4
episode: 689  score: 4.0  memory length: 128443  epsilon: 0.94368088000
12226  steps: 276  lr: 0.0001  evaluation reward: 1.43
episode: 690  score: 2.0  memory length: 128643  epsilon: 0.94328488000
12312  steps: 200  lr: 0.0001  evaluation reward: 1.44
episode: 691  score: 1.0  memory length: 128794  epsilon: 0.94298590000
12377  steps: 151  lr: 0.0001  evaluation reward: 1.45
episode: 692  score: 1.0  memory length: 128964  epsilon: 0.94264930000
1245  steps: 170  lr: 0.0001  evaluation reward: 1.44
episode: 693  score: 5.0  memory length: 129285  epsilon: 0.94201372000
12588  steps: 321  lr: 0.0001  evaluation reward: 1.46
episode: 694  score: 4.0  memory length: 129587  epsilon: 0.94141576000
12718  steps: 302  lr: 0.0001  evaluation reward: 1.49
episode: 695  score: 0.0  memory length: 129710  epsilon: 0.94117222000
12771  steps: 123  lr: 0.0001  evaluation reward: 1.49
episode: 696  score: 1.0  memory length: 129861  epsilon: 0.94087324000
12836  steps: 151  lr: 0.0001  evaluation reward: 1.5
episode: 697  score: 0.0  memory length: 129984  epsilon: 0.94062970000
12889  steps: 123  lr: 0.0001  evaluation reward: 1.5
episode: 698  score: 0.0  memory length: 130107  epsilon: 0.94038616000
12942  steps: 123  lr: 0.0001  evaluation reward: 1.46
episode: 699  score: 1.0  memory length: 130277  epsilon: 0.94004956000
13015  steps: 170  lr: 0.0001  evaluation reward: 1.43
episode: 700  score: 0.0  memory length: 130400  epsilon: 0.93980602000
13068  steps: 123  lr: 0.0001  evaluation reward: 1.42
episode: 701  score: 1.0  memory length: 130551  epsilon: 0.93950704000
13132  steps: 151  lr: 0.0001  evaluation reward: 1.43
episode: 702  score: 3.0  memory length: 130795  epsilon: 0.93902392000
13237  steps: 244  lr: 0.0001  evaluation reward: 1.44
episode: 703  score: 2.0  memory length: 131014  epsilon: 0.93859030000
13331  steps: 219  lr: 0.0001  evaluation reward: 1.46
episode: 704  score: 2.0  memory length: 131212  epsilon: 0.93819826000
13417  steps: 198  lr: 0.0001  evaluation reward: 1.47
episode: 705  score: 1.0  memory length: 131363  epsilon: 0.93789928000
13481  steps: 151  lr: 0.0001  evaluation reward: 1.45
episode: 706  score: 3.0  memory length: 131608  epsilon: 0.93741418000
13587  steps: 245  lr: 0.0001  evaluation reward: 1.47
episode: 707  score: 2.0  memory length: 131824  epsilon: 0.93698650000
1368  steps: 216  lr: 0.0001  evaluation reward: 1.48
episode: 708  score: 1.0  memory length: 131975  epsilon: 0.93668752000
13745  steps: 151  lr: 0.0001  evaluation reward: 1.46
episode: 709  score: 3.0  memory length: 132219  epsilon: 0.93620440000
13849  steps: 244  lr: 0.0001  evaluation reward: 1.47
episode: 710  score: 2.0  memory length: 132422  epsilon: 0.93580246000
13937  steps: 203  lr: 0.0001  evaluation reward: 1.46
```



```
episode: 711  score: 4.0  memory length: 132673  epsilon: 0.93530548000
14045  steps: 251  lr: 0.0001  evaluation reward: 1.49
episode: 712  score: 1.0  memory length: 132844  epsilon: 0.93496690000
14118  steps: 171  lr: 0.0001  evaluation reward: 1.49
episode: 713  score: 1.0  memory length: 133014  epsilon: 0.93463030000
14191  steps: 170  lr: 0.0001  evaluation reward: 1.48
episode: 714  score: 3.0  memory length: 133242  epsilon: 0.93417886000
14289  steps: 228  lr: 0.0001  evaluation reward: 1.49
episode: 715  score: 0.0  memory length: 133365  epsilon: 0.93393532000
14342  steps: 123  lr: 0.0001  evaluation reward: 1.49
episode: 716  score: 4.0  memory length: 133661  epsilon: 0.93334924000
14469  steps: 296  lr: 0.0001  evaluation reward: 1.48
episode: 717  score: 1.0  memory length: 133829  epsilon: 0.93301660000
14541  steps: 168  lr: 0.0001  evaluation reward: 1.49
episode: 718  score: 3.0  memory length: 134095  epsilon: 0.93248992000
14656  steps: 266  lr: 0.0001  evaluation reward: 1.5
episode: 719  score: 0.0  memory length: 134217  epsilon: 0.93224836000
14708  steps: 122  lr: 0.0001  evaluation reward: 1.5
episode: 720  score: 6.0  memory length: 134611  epsilon: 0.93146824000
14878  steps: 394  lr: 0.0001  evaluation reward: 1.54
episode: 721  score: 1.0  memory length: 134782  epsilon: 0.93112966000
14951  steps: 171  lr: 0.0001  evaluation reward: 1.54
episode: 722  score: 3.0  memory length: 135008  epsilon: 0.93068218000
15048  steps: 226  lr: 0.0001  evaluation reward: 1.57
episode: 723  score: 3.0  memory length: 135275  epsilon: 0.93015352000
15163  steps: 267  lr: 0.0001  evaluation reward: 1.59
episode: 724  score: 5.0  memory length: 135619  epsilon: 0.92947240000
15311  steps: 344  lr: 0.0001  evaluation reward: 1.64
episode: 725  score: 3.0  memory length: 135845  epsilon: 0.92902492000
15408  steps: 226  lr: 0.0001  evaluation reward: 1.66
episode: 726  score: 6.0  memory length: 136218  epsilon: 0.92828638000
15568  steps: 373  lr: 0.0001  evaluation reward: 1.71
episode: 727  score: 2.0  memory length: 136399  epsilon: 0.92792800000
15646  steps: 181  lr: 0.0001  evaluation reward: 1.73
episode: 728  score: 1.0  memory length: 136550  epsilon: 0.92762902000
15711  steps: 151  lr: 0.0001  evaluation reward: 1.74
episode: 729  score: 0.0  memory length: 136673  epsilon: 0.92738548000
15764  steps: 123  lr: 0.0001  evaluation reward: 1.74
episode: 730  score: 2.0  memory length: 136871  epsilon: 0.92699344000
15849  steps: 198  lr: 0.0001  evaluation reward: 1.74
episode: 731  score: 1.0  memory length: 137021  epsilon: 0.92669644000
15913  steps: 150  lr: 0.0001  evaluation reward: 1.75
episode: 732  score: 3.0  memory length: 137267  epsilon: 0.92620936000
16019  steps: 246  lr: 0.0001  evaluation reward: 1.78
episode: 733  score: 0.0  memory length: 137390  epsilon: 0.92596582000
16072  steps: 123  lr: 0.0001  evaluation reward: 1.77
episode: 734  score: 2.0  memory length: 137609  epsilon: 0.92553220000
16166  steps: 219  lr: 0.0001  evaluation reward: 1.77
episode: 735  score: 2.0  memory length: 137809  epsilon: 0.92513620000
16252  steps: 200  lr: 0.0001  evaluation reward: 1.79
episode: 736  score: 2.0  memory length: 137991  epsilon: 0.92477584000
1633  steps: 182  lr: 0.0001  evaluation reward: 1.81
episode: 737  score: 2.0  memory length: 138189  epsilon: 0.92438380000
```

```
16416  steps: 198  lr: 0.0001  evaluation reward: 1.83
episode: 738  score: 1.0  memory length: 138358  epsilon: 0.92404918000
16488  steps: 169  lr: 0.0001  evaluation reward: 1.82
episode: 739  score: 1.0  memory length: 138509  epsilon: 0.92375020000
16553  steps: 151  lr: 0.0001  evaluation reward: 1.76
episode: 740  score: 3.0  memory length: 138754  epsilon: 0.92326510000
16658  steps: 245  lr: 0.0001  evaluation reward: 1.79
episode: 741  score: 1.0  memory length: 138904  epsilon: 0.92296810000
16723  steps: 150  lr: 0.0001  evaluation reward: 1.76
episode: 742  score: 1.0  memory length: 139076  epsilon: 0.92262754000
16797  steps: 172  lr: 0.0001  evaluation reward: 1.77
episode: 743  score: 0.0  memory length: 139198  epsilon: 0.92238598000
16849  steps: 122  lr: 0.0001  evaluation reward: 1.77
episode: 744  score: 5.0  memory length: 139544  epsilon: 0.92170090000
16998  steps: 346  lr: 0.0001  evaluation reward: 1.82
episode: 745  score: 2.0  memory length: 139745  epsilon: 0.92130292000
17084  steps: 201  lr: 0.0001  evaluation reward: 1.82
episode: 746  score: 0.0  memory length: 139868  epsilon: 0.92105938000
17137  steps: 123  lr: 0.0001  evaluation reward: 1.8
episode: 747  score: 1.0  memory length: 140019  epsilon: 0.92076040000
17202  steps: 151  lr: 0.0001  evaluation reward: 1.79
episode: 748  score: 2.0  memory length: 140237  epsilon: 0.92032876000
17296  steps: 218  lr: 0.0001  evaluation reward: 1.81
episode: 749  score: 0.0  memory length: 140360  epsilon: 0.92008522000
17349  steps: 123  lr: 0.0001  evaluation reward: 1.79
episode: 750  score: 4.0  memory length: 140639  epsilon: 0.91953280000
17469  steps: 279  lr: 0.0001  evaluation reward: 1.83
episode: 751  score: 3.0  memory length: 140887  epsilon: 0.91904176000
17575  steps: 248  lr: 0.0001  evaluation reward: 1.86
episode: 752  score: 2.0  memory length: 141085  epsilon: 0.91864972000
1766  steps: 198  lr: 0.0001  evaluation reward: 1.88
episode: 753  score: 0.0  memory length: 141208  epsilon: 0.91840618000
17713  steps: 123  lr: 0.0001  evaluation reward: 1.88
episode: 754  score: 3.0  memory length: 141475  epsilon: 0.91787752000
17828  steps: 267  lr: 0.0001  evaluation reward: 1.91
episode: 755  score: 0.0  memory length: 141597  epsilon: 0.91763596000
1788  steps: 122  lr: 0.0001  evaluation reward: 1.91
episode: 756  score: 1.0  memory length: 141748  epsilon: 0.91733698000
17945  steps: 151  lr: 0.0001  evaluation reward: 1.89
episode: 757  score: 0.0  memory length: 141871  epsilon: 0.91709344000
17998  steps: 123  lr: 0.0001  evaluation reward: 1.87
episode: 758  score: 1.0  memory length: 142040  epsilon: 0.91675882000
18071  steps: 169  lr: 0.0001  evaluation reward: 1.87
episode: 759  score: 0.0  memory length: 142163  epsilon: 0.91651528000
18124  steps: 123  lr: 0.0001  evaluation reward: 1.85
episode: 760  score: 0.0  memory length: 142285  epsilon: 0.91627372000
18176  steps: 122  lr: 0.0001  evaluation reward: 1.85
episode: 761  score: 1.0  memory length: 142436  epsilon: 0.91597474000
18241  steps: 151  lr: 0.0001  evaluation reward: 1.84
episode: 762  score: 1.0  memory length: 142586  epsilon: 0.91567774000
18306  steps: 150  lr: 0.0001  evaluation reward: 1.82
episode: 763  score: 1.0  memory length: 142755  epsilon: 0.91534312000
18378  steps: 169  lr: 0.0001  evaluation reward: 1.81
```

episode: 764 score: 2.0 memory length: 142952 epsilon: 0.91495306000
18463 steps: 197 lr: 0.0001 evaluation reward: 1.82
episode: 765 score: 1.0 memory length: 143103 epsilon: 0.91465408000
18528 steps: 151 lr: 0.0001 evaluation reward: 1.83
episode: 766 score: 2.0 memory length: 143301 epsilon: 0.91426204000
18613 steps: 198 lr: 0.0001 evaluation reward: 1.85
episode: 767 score: 1.0 memory length: 143469 epsilon: 0.91392940000
18685 steps: 168 lr: 0.0001 evaluation reward: 1.83
episode: 768 score: 4.0 memory length: 143728 epsilon: 0.91341658000
18796 steps: 259 lr: 0.0001 evaluation reward: 1.85
episode: 769 score: 2.0 memory length: 143908 epsilon: 0.91306018000
18874 steps: 180 lr: 0.0001 evaluation reward: 1.83
episode: 770 score: 3.0 memory length: 144176 epsilon: 0.91252954000
18989 steps: 268 lr: 0.0001 evaluation reward: 1.86
episode: 771 score: 1.0 memory length: 144345 epsilon: 0.91219492000
19062 steps: 169 lr: 0.0001 evaluation reward: 1.84
episode: 772 score: 3.0 memory length: 144574 epsilon: 0.91174150000
1916 steps: 229 lr: 0.0001 evaluation reward: 1.85
episode: 773 score: 3.0 memory length: 144824 epsilon: 0.91124650000
19268 steps: 250 lr: 0.0001 evaluation reward: 1.86
episode: 774 score: 2.0 memory length: 145022 epsilon: 0.91085446000
19353 steps: 198 lr: 0.0001 evaluation reward: 1.86
episode: 775 score: 4.0 memory length: 145282 epsilon: 0.91033966000
19464 steps: 260 lr: 0.0001 evaluation reward: 1.86
episode: 776 score: 1.0 memory length: 145452 epsilon: 0.91000306000
19537 steps: 170 lr: 0.0001 evaluation reward: 1.84
episode: 777 score: 3.0 memory length: 145679 epsilon: 0.90955360000
19635 steps: 227 lr: 0.0001 evaluation reward: 1.87
episode: 778 score: 4.0 memory length: 145953 epsilon: 0.90901108000
19753 steps: 274 lr: 0.0001 evaluation reward: 1.9
episode: 779 score: 2.0 memory length: 146153 epsilon: 0.90861508000
19839 steps: 200 lr: 0.0001 evaluation reward: 1.9
episode: 780 score: 0.0 memory length: 146276 epsilon: 0.90837154000
19892 steps: 123 lr: 0.0001 evaluation reward: 1.9
episode: 781 score: 0.0 memory length: 146398 epsilon: 0.90812998000
19944 steps: 122 lr: 0.0001 evaluation reward: 1.85
episode: 782 score: 1.0 memory length: 146549 epsilon: 0.90783100000
20009 steps: 151 lr: 0.0001 evaluation reward: 1.83
episode: 783 score: 1.0 memory length: 146721 epsilon: 0.90749044000
20083 steps: 172 lr: 0.0001 evaluation reward: 1.81
episode: 784 score: 1.0 memory length: 146891 epsilon: 0.90715384000
20156 steps: 170 lr: 0.0001 evaluation reward: 1.81
episode: 785 score: 1.0 memory length: 147042 epsilon: 0.90685486000
20221 steps: 151 lr: 0.0001 evaluation reward: 1.79
episode: 786 score: 1.0 memory length: 147193 epsilon: 0.90655588000
20286 steps: 151 lr: 0.0001 evaluation reward: 1.79
episode: 787 score: 2.0 memory length: 147391 epsilon: 0.90616384000
20371 steps: 198 lr: 0.0001 evaluation reward: 1.81
episode: 788 score: 4.0 memory length: 147705 epsilon: 0.90554212000
20506 steps: 314 lr: 0.0001 evaluation reward: 1.83
episode: 789 score: 1.0 memory length: 147873 epsilon: 0.90520948000
20578 steps: 168 lr: 0.0001 evaluation reward: 1.8
episode: 790 score: 3.0 memory length: 148099 epsilon: 0.90476200000

```
20675  steps: 226  lr: 0.0001  evaluation reward: 1.81
episode: 791  score: 1.0  memory length: 148267  epsilon: 0.90442936000
20747  steps: 168  lr: 0.0001  evaluation reward: 1.81
episode: 792  score: 7.0  memory length: 148661  epsilon: 0.90364924000
20917  steps: 394  lr: 0.0001  evaluation reward: 1.87
episode: 793  score: 2.0  memory length: 148879  epsilon: 0.90321760000
2101  steps: 218  lr: 0.0001  evaluation reward: 1.84
episode: 794  score: 0.0  memory length: 149002  epsilon: 0.90297406000
21063  steps: 123  lr: 0.0001  evaluation reward: 1.8
episode: 795  score: 3.0  memory length: 149229  epsilon: 0.90252460000
21161  steps: 227  lr: 0.0001  evaluation reward: 1.83
episode: 796  score: 0.0  memory length: 149352  epsilon: 0.90228106000
21214  steps: 123  lr: 0.0001  evaluation reward: 1.82
episode: 797  score: 1.0  memory length: 149522  epsilon: 0.90194446000
21287  steps: 170  lr: 0.0001  evaluation reward: 1.83
episode: 798  score: 4.0  memory length: 149816  epsilon: 0.90136234000
21413  steps: 294  lr: 0.0001  evaluation reward: 1.87
episode: 799  score: 3.0  memory length: 150064  epsilon: 0.90087130000
2152  steps: 248  lr: 0.0001  evaluation reward: 1.89
episode: 800  score: 3.0  memory length: 150310  epsilon: 0.90038422000
21626  steps: 246  lr: 0.0001  evaluation reward: 1.92
episode: 801  score: 3.0  memory length: 150556  epsilon: 0.89989714000
21731  steps: 246  lr: 0.0001  evaluation reward: 1.94
episode: 802  score: 2.0  memory length: 150774  epsilon: 0.89946550000
21825  steps: 218  lr: 0.0001  evaluation reward: 1.93
episode: 803  score: 2.0  memory length: 150974  epsilon: 0.89906950000
21911  steps: 200  lr: 0.0001  evaluation reward: 1.93
episode: 804  score: 2.0  memory length: 151174  epsilon: 0.89867350000
21997  steps: 200  lr: 0.0001  evaluation reward: 1.93
episode: 805  score: 1.0  memory length: 151325  epsilon: 0.89837452000
22062  steps: 151  lr: 0.0001  evaluation reward: 1.93
episode: 806  score: 0.0  memory length: 151448  epsilon: 0.89813098000
22115  steps: 123  lr: 0.0001  evaluation reward: 1.9
episode: 807  score: 3.0  memory length: 151695  epsilon: 0.89764192000
22221  steps: 247  lr: 0.0001  evaluation reward: 1.91
episode: 808  score: 5.0  memory length: 152003  epsilon: 0.89703208000
22353  steps: 308  lr: 0.0001  evaluation reward: 1.95
episode: 809  score: 3.0  memory length: 152232  epsilon: 0.89657866000
22452  steps: 229  lr: 0.0001  evaluation reward: 1.95
episode: 810  score: 1.0  memory length: 152403  epsilon: 0.89624008000
22525  steps: 171  lr: 0.0001  evaluation reward: 1.94
episode: 811  score: 4.0  memory length: 152698  epsilon: 0.89565598000
22652  steps: 295  lr: 0.0001  evaluation reward: 1.94
episode: 812  score: 4.0  memory length: 152993  epsilon: 0.89507188000
22779  steps: 295  lr: 0.0001  evaluation reward: 1.97
episode: 813  score: 4.0  memory length: 153311  epsilon: 0.89444224000
22916  steps: 318  lr: 0.0001  evaluation reward: 2.0
episode: 814  score: 0.0  memory length: 153434  epsilon: 0.89419870000
22968  steps: 123  lr: 0.0001  evaluation reward: 1.97
episode: 815  score: 4.0  memory length: 153713  epsilon: 0.89364628000
23088  steps: 279  lr: 0.0001  evaluation reward: 2.01
episode: 816  score: 2.0  memory length: 153910  epsilon: 0.89325622000
23173  steps: 197  lr: 0.0001  evaluation reward: 1.99
```

```
episode: 817  score: 4.0  memory length: 154227  epsilon: 0.89262856000
23309  steps: 317  lr: 0.0001  evaluation reward: 2.02
episode: 818  score: 0.0  memory length: 154349  epsilon: 0.89238700000
23362  steps: 122  lr: 0.0001  evaluation reward: 1.99
episode: 819  score: 2.0  memory length: 154547  epsilon: 0.89199496000
23447  steps: 198  lr: 0.0001  evaluation reward: 2.01
episode: 820  score: 8.0  memory length: 154877  epsilon: 0.89134156000
23589  steps: 330  lr: 0.0001  evaluation reward: 2.03
episode: 821  score: 3.0  memory length: 155103  epsilon: 0.89089408000
23686  steps: 226  lr: 0.0001  evaluation reward: 2.05
episode: 822  score: 2.0  memory length: 155300  epsilon: 0.89050402000
2377  steps: 197  lr: 0.0001  evaluation reward: 2.04
episode: 823  score: 6.0  memory length: 155642  epsilon: 0.88982686000
23918  steps: 342  lr: 0.0001  evaluation reward: 2.07
episode: 824  score: 1.0  memory length: 155810  epsilon: 0.88949422000
2399  steps: 168  lr: 0.0001  evaluation reward: 2.03
episode: 825  score: 1.0  memory length: 155979  epsilon: 0.88915960000
24062  steps: 169  lr: 0.0001  evaluation reward: 2.01
episode: 826  score: 1.0  memory length: 156131  epsilon: 0.88885864000
24128  steps: 152  lr: 0.0001  evaluation reward: 1.96
episode: 827  score: 3.0  memory length: 156395  epsilon: 0.88833592000
24241  steps: 264  lr: 0.0001  evaluation reward: 1.97
episode: 828  score: 1.0  memory length: 156546  epsilon: 0.88803694000
24306  steps: 151  lr: 0.0001  evaluation reward: 1.97
episode: 829  score: 0.0  memory length: 156669  epsilon: 0.88779340000
24359  steps: 123  lr: 0.0001  evaluation reward: 1.97
episode: 830  score: 5.0  memory length: 156994  epsilon: 0.88714990000
24499  steps: 325  lr: 0.0001  evaluation reward: 2.0
episode: 831  score: 3.0  memory length: 157238  epsilon: 0.88666678000
24604  steps: 244  lr: 0.0001  evaluation reward: 2.02
episode: 832  score: 0.0  memory length: 157361  epsilon: 0.88642324000
24656  steps: 123  lr: 0.0001  evaluation reward: 1.99
episode: 833  score: 0.0  memory length: 157483  epsilon: 0.88618168000
24709  steps: 122  lr: 0.0001  evaluation reward: 1.99
episode: 834  score: 2.0  memory length: 157663  epsilon: 0.88582528000
24786  steps: 180  lr: 0.0001  evaluation reward: 1.99
episode: 835  score: 4.0  memory length: 157938  epsilon: 0.88528078000
24904  steps: 275  lr: 0.0001  evaluation reward: 2.01
episode: 836  score: 1.0  memory length: 158088  epsilon: 0.88498378000
24969  steps: 150  lr: 0.0001  evaluation reward: 2.0
episode: 837  score: 3.0  memory length: 158359  epsilon: 0.88444720000
25085  steps: 271  lr: 0.0001  evaluation reward: 2.01
episode: 838  score: 3.0  memory length: 158607  epsilon: 0.88395616000
25192  steps: 248  lr: 0.0001  evaluation reward: 2.03
episode: 839  score: 3.0  memory length: 158833  epsilon: 0.88350868000
25289  steps: 226  lr: 0.0001  evaluation reward: 2.05
episode: 840  score: 3.0  memory length: 159098  epsilon: 0.88298398000
25403  steps: 265  lr: 0.0001  evaluation reward: 2.05
episode: 841  score: 2.0  memory length: 159278  epsilon: 0.88262758000
2548  steps: 180  lr: 0.0001  evaluation reward: 2.06
episode: 842  score: 8.0  memory length: 159624  epsilon: 0.88194250000
25629  steps: 346  lr: 0.0001  evaluation reward: 2.13
episode: 843  score: 1.0  memory length: 159794  epsilon: 0.88160590000
```

```
25702  steps: 170    lr: 0.0001    evaluation reward: 2.14
episode: 844  score: 2.0    memory length: 159991    epsilon: 0.88121584000
25787  steps: 197    lr: 0.0001    evaluation reward: 2.11
episode: 845  score: 1.0    memory length: 160141    epsilon: 0.88091884000
25851  steps: 150    lr: 0.0001    evaluation reward: 2.1
episode: 846  score: 3.0    memory length: 160387    epsilon: 0.88043176000
25957  steps: 246    lr: 0.0001    evaluation reward: 2.13
episode: 847  score: 2.0    memory length: 160585    epsilon: 0.88003972000
26042  steps: 198    lr: 0.0001    evaluation reward: 2.14
episode: 848  score: 2.0    memory length: 160782    epsilon: 0.87964966000
26127  steps: 197    lr: 0.0001    evaluation reward: 2.14
episode: 849  score: 2.0    memory length: 160981    epsilon: 0.87925564000
26212  steps: 199    lr: 0.0001    evaluation reward: 2.16
episode: 850  score: 4.0    memory length: 161243    epsilon: 0.87873688000
26325  steps: 262    lr: 0.0001    evaluation reward: 2.16
episode: 851  score: 4.0    memory length: 161533    epsilon: 0.87816268000
2645   steps: 290    lr: 0.0001    evaluation reward: 2.17
episode: 852  score: 3.0    memory length: 161799    epsilon: 0.87763600000
26564  steps: 266    lr: 0.0001    evaluation reward: 2.18
episode: 853  score: 3.0    memory length: 162025    epsilon: 0.87718852000
26661  steps: 226    lr: 0.0001    evaluation reward: 2.21
episode: 854  score: 4.0    memory length: 162301    epsilon: 0.87664204000
2678   steps: 276    lr: 0.0001    evaluation reward: 2.22
episode: 855  score: 4.0    memory length: 162598    epsilon: 0.87605398000
26907  steps: 297    lr: 0.0001    evaluation reward: 2.26
episode: 856  score: 2.0    memory length: 162795    epsilon: 0.87566392000
26992  steps: 197    lr: 0.0001    evaluation reward: 2.27
episode: 857  score: 3.0    memory length: 163041    epsilon: 0.87517684000
27098  steps: 246    lr: 0.0001    evaluation reward: 2.3
episode: 858  score: 1.0    memory length: 163212    epsilon: 0.87483826000
27171  steps: 171    lr: 0.0001    evaluation reward: 2.3
episode: 859  score: 3.0    memory length: 163459    epsilon: 0.87434920000
27278  steps: 247    lr: 0.0001    evaluation reward: 2.33
episode: 860  score: 4.0    memory length: 163757    epsilon: 0.87375916000
27406  steps: 298    lr: 0.0001    evaluation reward: 2.37
episode: 861  score: 2.0    memory length: 163974    epsilon: 0.87332950000
27499  steps: 217    lr: 0.0001    evaluation reward: 2.38
episode: 862  score: 2.0    memory length: 164154    epsilon: 0.87297310000
27576  steps: 180    lr: 0.0001    evaluation reward: 2.39
episode: 863  score: 2.0    memory length: 164356    epsilon: 0.87257314000
27663  steps: 202    lr: 0.0001    evaluation reward: 2.4
episode: 864  score: 5.0    memory length: 164673    epsilon: 0.87194548000
27799  steps: 317    lr: 0.0001    evaluation reward: 2.43
episode: 865  score: 4.0    memory length: 164947    epsilon: 0.87140296000
27917  steps: 274    lr: 0.0001    evaluation reward: 2.46
episode: 866  score: 1.0    memory length: 165117    epsilon: 0.87106636000
2799   steps: 170    lr: 0.0001    evaluation reward: 2.45
episode: 867  score: 2.0    memory length: 165335    epsilon: 0.87063472000
28084  steps: 218    lr: 0.0001    evaluation reward: 2.46
episode: 868  score: 2.0    memory length: 165532    epsilon: 0.87024466000
28169  steps: 197    lr: 0.0001    evaluation reward: 2.44
episode: 869  score: 6.0    memory length: 165869    epsilon: 0.86957740000
28313  steps: 337    lr: 0.0001    evaluation reward: 2.48
```

episode: 870 score: 5.0 memory length: 166194 epsilon: 0.86893390000
28453 steps: 325 lr: 0.0001 evaluation reward: 2.5
episode: 871 score: 5.0 memory length: 166520 epsilon: 0.86828842000
28593 steps: 326 lr: 0.0001 evaluation reward: 2.54
episode: 872 score: 3.0 memory length: 166786 epsilon: 0.86776174000
28708 steps: 266 lr: 0.0001 evaluation reward: 2.54
episode: 873 score: 5.0 memory length: 167113 epsilon: 0.86711428000
28848 steps: 327 lr: 0.0001 evaluation reward: 2.56
episode: 874 score: 11.0 memory length: 167545 epsilon: 0.8662589200
029034 steps: 432 lr: 0.0001 evaluation reward: 2.65
episode: 875 score: 3.0 memory length: 167790 epsilon: 0.86577382000
29139 steps: 245 lr: 0.0001 evaluation reward: 2.64
episode: 876 score: 3.0 memory length: 168017 epsilon: 0.86532436000
29237 steps: 227 lr: 0.0001 evaluation reward: 2.66
episode: 877 score: 2.0 memory length: 168218 epsilon: 0.86492638000
29323 steps: 201 lr: 0.0001 evaluation reward: 2.65
episode: 878 score: 3.0 memory length: 168465 epsilon: 0.86443732000
29429 steps: 247 lr: 0.0001 evaluation reward: 2.64
episode: 879 score: 1.0 memory length: 168616 epsilon: 0.86413834000
29494 steps: 151 lr: 0.0001 evaluation reward: 2.63
episode: 880 score: 6.0 memory length: 168959 epsilon: 0.86345920000
29642 steps: 343 lr: 0.0001 evaluation reward: 2.69
episode: 881 score: 1.0 memory length: 169110 epsilon: 0.86316022000
29707 steps: 151 lr: 0.0001 evaluation reward: 2.7
episode: 882 score: 2.0 memory length: 169308 epsilon: 0.86276818000
29792 steps: 198 lr: 0.0001 evaluation reward: 2.71
episode: 883 score: 0.0 memory length: 169431 epsilon: 0.86252464000
29845 steps: 123 lr: 0.0001 evaluation reward: 2.7
episode: 884 score: 3.0 memory length: 169676 epsilon: 0.86203954000
2995 steps: 245 lr: 0.0001 evaluation reward: 2.72
episode: 885 score: 5.0 memory length: 169972 epsilon: 0.86145346000
30077 steps: 296 lr: 0.0001 evaluation reward: 2.76
episode: 886 score: 5.0 memory length: 170287 epsilon: 0.86082976000
30212 steps: 315 lr: 0.0001 evaluation reward: 2.8
episode: 887 score: 3.0 memory length: 170534 epsilon: 0.86034070000
30319 steps: 247 lr: 0.0001 evaluation reward: 2.81
episode: 888 score: 1.0 memory length: 170685 epsilon: 0.86004172000
30384 steps: 151 lr: 0.0001 evaluation reward: 2.78
episode: 889 score: 6.0 memory length: 171058 epsilon: 0.85930318000
30544 steps: 373 lr: 0.0001 evaluation reward: 2.83
episode: 890 score: 3.0 memory length: 171286 epsilon: 0.85885174000
30642 steps: 228 lr: 0.0001 evaluation reward: 2.83
episode: 891 score: 2.0 memory length: 171484 epsilon: 0.85845970000
30727 steps: 198 lr: 0.0001 evaluation reward: 2.84
episode: 892 score: 1.0 memory length: 171635 epsilon: 0.85816072000
30792 steps: 151 lr: 0.0001 evaluation reward: 2.78
episode: 893 score: 2.0 memory length: 171833 epsilon: 0.85776868000
30877 steps: 198 lr: 0.0001 evaluation reward: 2.78
episode: 894 score: 2.0 memory length: 172051 epsilon: 0.85733704000
30971 steps: 218 lr: 0.0001 evaluation reward: 2.8
episode: 895 score: 5.0 memory length: 172397 epsilon: 0.85665196000
3112 steps: 346 lr: 0.0001 evaluation reward: 2.82
episode: 896 score: 3.0 memory length: 172663 epsilon: 0.85612528000

```
31234 steps: 266 lr: 0.0001 evaluation reward: 2.85
episode: 897 score: 4.0 memory length: 172961 epsilon: 0.85553524000
31362 steps: 298 lr: 0.0001 evaluation reward: 2.88
episode: 898 score: 2.0 memory length: 173162 epsilon: 0.85513726000
31448 steps: 201 lr: 0.0001 evaluation reward: 2.86
episode: 899 score: 4.0 memory length: 173420 epsilon: 0.85462642000
31559 steps: 258 lr: 0.0001 evaluation reward: 2.87
episode: 900 score: 0.0 memory length: 173542 epsilon: 0.85438486000
31612 steps: 122 lr: 0.0001 evaluation reward: 2.84
episode: 901 score: 3.0 memory length: 173768 epsilon: 0.85393738000
31709 steps: 226 lr: 0.0001 evaluation reward: 2.84
episode: 902 score: 2.0 memory length: 173990 epsilon: 0.85349782000
31804 steps: 222 lr: 0.0001 evaluation reward: 2.84
episode: 903 score: 2.0 memory length: 174170 epsilon: 0.85314142000
31882 steps: 180 lr: 0.0001 evaluation reward: 2.84
episode: 904 score: 2.0 memory length: 174350 epsilon: 0.85278502000
31959 steps: 180 lr: 0.0001 evaluation reward: 2.84
episode: 905 score: 3.0 memory length: 174579 epsilon: 0.85233160000
32057 steps: 229 lr: 0.0001 evaluation reward: 2.86
episode: 906 score: 3.0 memory length: 174847 epsilon: 0.85180096000
32173 steps: 268 lr: 0.0001 evaluation reward: 2.89
episode: 907 score: 5.0 memory length: 175156 epsilon: 0.85118914000
32305 steps: 309 lr: 0.0001 evaluation reward: 2.91
episode: 908 score: 5.0 memory length: 175478 epsilon: 0.85055158000
32444 steps: 322 lr: 0.0001 evaluation reward: 2.91
episode: 909 score: 4.0 memory length: 175755 epsilon: 0.85000312000
32563 steps: 277 lr: 0.0001 evaluation reward: 2.92
episode: 910 score: 4.0 memory length: 176031 epsilon: 0.84945664000
32681 steps: 276 lr: 0.0001 evaluation reward: 2.95
episode: 911 score: 3.0 memory length: 176241 epsilon: 0.84904084000
32772 steps: 210 lr: 0.0001 evaluation reward: 2.94
episode: 912 score: 3.0 memory length: 176492 epsilon: 0.84854386000
3288 steps: 251 lr: 0.0001 evaluation reward: 2.93
episode: 913 score: 2.0 memory length: 176690 epsilon: 0.84815182000
32965 steps: 198 lr: 0.0001 evaluation reward: 2.91
episode: 914 score: 0.0 memory length: 176813 epsilon: 0.84790828000
33018 steps: 123 lr: 0.0001 evaluation reward: 2.91
episode: 915 score: 4.0 memory length: 177086 epsilon: 0.84736774000
33135 steps: 273 lr: 0.0001 evaluation reward: 2.91
episode: 916 score: 4.0 memory length: 177365 epsilon: 0.84681532000
33255 steps: 279 lr: 0.0001 evaluation reward: 2.93
episode: 917 score: 4.0 memory length: 177661 epsilon: 0.84622924000
33382 steps: 296 lr: 0.0001 evaluation reward: 2.93
episode: 918 score: 3.0 memory length: 177927 epsilon: 0.84570256000
33496 steps: 266 lr: 0.0001 evaluation reward: 2.96
episode: 919 score: 3.0 memory length: 178172 epsilon: 0.84521746000
33602 steps: 245 lr: 0.0001 evaluation reward: 2.97
episode: 920 score: 2.0 memory length: 178370 epsilon: 0.84482542000
33687 steps: 198 lr: 0.0001 evaluation reward: 2.91
episode: 921 score: 0.0 memory length: 178493 epsilon: 0.84458188000
3374 steps: 123 lr: 0.0001 evaluation reward: 2.88
episode: 922 score: 3.0 memory length: 178737 epsilon: 0.84409876000
33845 steps: 244 lr: 0.0001 evaluation reward: 2.89
```



```
episode: 923  score: 2.0  memory length: 178935  epsilon: 0.84370672000
3393  steps: 198  lr: 0.0001  evaluation reward: 2.85
episode: 924  score: 1.0  memory length: 179104  epsilon: 0.84337210000
34002  steps: 169  lr: 0.0001  evaluation reward: 2.85
episode: 925  score: 4.0  memory length: 179400  epsilon: 0.84278602000
3413  steps: 296  lr: 0.0001  evaluation reward: 2.88
episode: 926  score: 3.0  memory length: 179643  epsilon: 0.84230488000
34234  steps: 243  lr: 0.0001  evaluation reward: 2.9
episode: 927  score: 1.0  memory length: 179794  epsilon: 0.84200590000
34299  steps: 151  lr: 0.0001  evaluation reward: 2.88
episode: 928  score: 3.0  memory length: 180019  epsilon: 0.84156040000
34396  steps: 225  lr: 0.0001  evaluation reward: 2.9
episode: 929  score: 1.0  memory length: 180188  epsilon: 0.84122578000
34468  steps: 169  lr: 0.0001  evaluation reward: 2.91
episode: 930  score: 1.0  memory length: 180359  epsilon: 0.84088720000
34542  steps: 171  lr: 0.0001  evaluation reward: 2.87
episode: 931  score: 0.0  memory length: 180482  epsilon: 0.84064366000
34595  steps: 123  lr: 0.0001  evaluation reward: 2.84
episode: 932  score: 0.0  memory length: 180604  epsilon: 0.84040210000
34647  steps: 122  lr: 0.0001  evaluation reward: 2.84
episode: 933  score: 6.0  memory length: 180978  epsilon: 0.83966158000
34808  steps: 374  lr: 0.0001  evaluation reward: 2.9
episode: 934  score: 2.0  memory length: 181179  epsilon: 0.83926360000
34894  steps: 201  lr: 0.0001  evaluation reward: 2.9
episode: 935  score: 2.0  memory length: 181377  epsilon: 0.83887156000
34979  steps: 198  lr: 0.0001  evaluation reward: 2.88
episode: 936  score: 3.0  memory length: 181605  epsilon: 0.83842012000
35077  steps: 228  lr: 0.0001  evaluation reward: 2.9
episode: 937  score: 3.0  memory length: 181832  epsilon: 0.83797066000
35175  steps: 227  lr: 0.0001  evaluation reward: 2.9
episode: 938  score: 4.0  memory length: 182107  epsilon: 0.83742616000
35293  steps: 275  lr: 0.0001  evaluation reward: 2.91
episode: 939  score: 2.0  memory length: 182324  epsilon: 0.83699650000
35386  steps: 217  lr: 0.0001  evaluation reward: 2.9
episode: 940  score: 2.0  memory length: 182522  epsilon: 0.83660446000
35472  steps: 198  lr: 0.0001  evaluation reward: 2.89
episode: 941  score: 3.0  memory length: 182769  epsilon: 0.83611540000
35578  steps: 247  lr: 0.0001  evaluation reward: 2.9
episode: 942  score: 3.0  memory length: 183019  epsilon: 0.83562040000
35685  steps: 250  lr: 0.0001  evaluation reward: 2.85
episode: 943  score: 2.0  memory length: 183236  epsilon: 0.83519074000
35778  steps: 217  lr: 0.0001  evaluation reward: 2.86
episode: 944  score: 2.0  memory length: 183436  epsilon: 0.83479474000
35864  steps: 200  lr: 0.0001  evaluation reward: 2.86
episode: 945  score: 2.0  memory length: 183617  epsilon: 0.83443636000
35942  steps: 181  lr: 0.0001  evaluation reward: 2.87
episode: 946  score: 4.0  memory length: 183899  epsilon: 0.83387800000
36063  steps: 282  lr: 0.0001  evaluation reward: 2.88
episode: 947  score: 4.0  memory length: 184175  epsilon: 0.83333152000
36182  steps: 276  lr: 0.0001  evaluation reward: 2.9
episode: 948  score: 3.0  memory length: 184420  epsilon: 0.83284642000
36287  steps: 245  lr: 0.0001  evaluation reward: 2.91
episode: 949  score: 2.0  memory length: 184619  epsilon: 0.83245240000
```

```
36373  steps: 199  lr: 0.0001  evaluation reward: 2.91
episode: 950  score: 2.0  memory length: 184818  epsilon: 0.83205838000
36458  steps: 199  lr: 0.0001  evaluation reward: 2.89
episode: 951  score: 5.0  memory length: 185141  epsilon: 0.83141884000
36597  steps: 323  lr: 0.0001  evaluation reward: 2.9
episode: 952  score: 0.0  memory length: 185264  epsilon: 0.83117530000
3665  steps: 123  lr: 0.0001  evaluation reward: 2.87
episode: 953  score: 0.0  memory length: 185387  epsilon: 0.83093176000
36703  steps: 123  lr: 0.0001  evaluation reward: 2.84
episode: 954  score: 4.0  memory length: 185683  epsilon: 0.83034568000
3683  steps: 296  lr: 0.0001  evaluation reward: 2.84
episode: 955  score: 5.0  memory length: 185990  epsilon: 0.82973782000
36962  steps: 307  lr: 0.0001  evaluation reward: 2.85
episode: 956  score: 5.0  memory length: 186313  epsilon: 0.82909828000
37101  steps: 323  lr: 0.0001  evaluation reward: 2.88
episode: 957  score: 0.0  memory length: 186435  epsilon: 0.82885672000
37154  steps: 122  lr: 0.0001  evaluation reward: 2.85
episode: 958  score: 2.0  memory length: 186652  epsilon: 0.82842706000
37247  steps: 217  lr: 0.0001  evaluation reward: 2.86
episode: 959  score: 4.0  memory length: 186927  epsilon: 0.82788256000
37365  steps: 275  lr: 0.0001  evaluation reward: 2.87
episode: 960  score: 0.0  memory length: 187050  epsilon: 0.82763902000
37418  steps: 123  lr: 0.0001  evaluation reward: 2.83
episode: 961  score: 0.0  memory length: 187172  epsilon: 0.82739746000
3747  steps: 122  lr: 0.0001  evaluation reward: 2.81
episode: 962  score: 3.0  memory length: 187437  epsilon: 0.82687276000
37584  steps: 265  lr: 0.0001  evaluation reward: 2.82
episode: 963  score: 5.0  memory length: 187748  epsilon: 0.82625698000
37718  steps: 311  lr: 0.0001  evaluation reward: 2.85
episode: 964  score: 3.0  memory length: 187974  epsilon: 0.82580950000
37815  steps: 226  lr: 0.0001  evaluation reward: 2.83
episode: 965  score: 0.0  memory length: 188097  epsilon: 0.82556596000
37868  steps: 123  lr: 0.0001  evaluation reward: 2.79
episode: 966  score: 1.0  memory length: 188248  epsilon: 0.82526698000
37933  steps: 151  lr: 0.0001  evaluation reward: 2.79
episode: 967  score: 4.0  memory length: 188549  epsilon: 0.82467100000
38062  steps: 301  lr: 0.0001  evaluation reward: 2.81
episode: 968  score: 1.0  memory length: 188718  epsilon: 0.82433638000
38135  steps: 169  lr: 0.0001  evaluation reward: 2.8
episode: 969  score: 3.0  memory length: 188946  epsilon: 0.82388494000
38233  steps: 228  lr: 0.0001  evaluation reward: 2.77
episode: 970  score: 2.0  memory length: 189163  epsilon: 0.82345528000
38326  steps: 217  lr: 0.0001  evaluation reward: 2.74
episode: 971  score: 3.0  memory length: 189390  epsilon: 0.82300582000
38424  steps: 227  lr: 0.0001  evaluation reward: 2.72
episode: 972  score: 4.0  memory length: 189648  epsilon: 0.82249498000
38535  steps: 258  lr: 0.0001  evaluation reward: 2.73
episode: 973  score: 4.0  memory length: 189945  epsilon: 0.82190692000
38662  steps: 297  lr: 0.0001  evaluation reward: 2.72
episode: 974  score: 2.0  memory length: 190163  epsilon: 0.82147528000
38756  steps: 218  lr: 0.0001  evaluation reward: 2.63
episode: 975  score: 3.0  memory length: 190409  epsilon: 0.82098820000
38862  steps: 246  lr: 0.0001  evaluation reward: 2.63
```

```
episode: 976  score: 1.0  memory length: 190560  epsilon: 0.82068922000
38927  steps: 151  lr: 0.0001  evaluation reward: 2.61
episode: 977  score: 0.0  memory length: 190682  epsilon: 0.82044766000
38979  steps: 122  lr: 0.0001  evaluation reward: 2.59
episode: 978  score: 1.0  memory length: 190852  epsilon: 0.82011106000
39052  steps: 170  lr: 0.0001  evaluation reward: 2.57
episode: 979  score: 2.0  memory length: 191050  epsilon: 0.81971902000
39137  steps: 198  lr: 0.0001  evaluation reward: 2.58
episode: 980  score: 3.0  memory length: 191296  epsilon: 0.81923194000
39243  steps: 246  lr: 0.0001  evaluation reward: 2.55
episode: 981  score: 0.0  memory length: 191418  epsilon: 0.81899038000
39295  steps: 122  lr: 0.0001  evaluation reward: 2.54
episode: 982  score: 2.0  memory length: 191602  epsilon: 0.81862606000
39375  steps: 184  lr: 0.0001  evaluation reward: 2.54
episode: 983  score: 3.0  memory length: 191849  epsilon: 0.81813700000
39481  steps: 247  lr: 0.0001  evaluation reward: 2.57
episode: 984  score: 1.0  memory length: 192018  epsilon: 0.81780238000
39553  steps: 169  lr: 0.0001  evaluation reward: 2.55
episode: 985  score: 0.0  memory length: 192141  epsilon: 0.81755884000
39606  steps: 123  lr: 0.0001  evaluation reward: 2.5
episode: 986  score: 4.0  memory length: 192397  epsilon: 0.81705196000
39716  steps: 256  lr: 0.0001  evaluation reward: 2.49
episode: 987  score: 2.0  memory length: 192597  epsilon: 0.81665596000
39802  steps: 200  lr: 0.0001  evaluation reward: 2.48
episode: 988  score: 1.0  memory length: 192748  epsilon: 0.81635698000
39867  steps: 151  lr: 0.0001  evaluation reward: 2.48
episode: 989  score: 2.0  memory length: 192945  epsilon: 0.81596692000
39952  steps: 197  lr: 0.0001  evaluation reward: 2.44
episode: 990  score: 0.0  memory length: 193068  epsilon: 0.81572338000
40005  steps: 123  lr: 0.0001  evaluation reward: 2.41
episode: 991  score: 6.0  memory length: 193320  epsilon: 0.81522442000
40113  steps: 252  lr: 0.0001  evaluation reward: 2.45
episode: 992  score: 2.0  memory length: 193520  epsilon: 0.81482842000
40199  steps: 200  lr: 0.0001  evaluation reward: 2.46
episode: 993  score: 2.0  memory length: 193717  epsilon: 0.81443836000
40284  steps: 197  lr: 0.0001  evaluation reward: 2.46
episode: 994  score: 3.0  memory length: 193967  epsilon: 0.81394336000
40391  steps: 250  lr: 0.0001  evaluation reward: 2.47
episode: 995  score: 3.0  memory length: 194214  epsilon: 0.81345430000
40497  steps: 247  lr: 0.0001  evaluation reward: 2.45
episode: 996  score: 4.0  memory length: 194511  epsilon: 0.81286624000
40625  steps: 297  lr: 0.0001  evaluation reward: 2.46
episode: 997  score: 3.0  memory length: 194758  epsilon: 0.81237718000
40731  steps: 247  lr: 0.0001  evaluation reward: 2.45
episode: 998  score: 5.0  memory length: 195085  epsilon: 0.81172972000
40872  steps: 327  lr: 0.0001  evaluation reward: 2.48
episode: 999  score: 2.0  memory length: 195283  epsilon: 0.81133768000
40957  steps: 198  lr: 0.0001  evaluation reward: 2.46
episode: 1000 score: 0.0  memory length: 195406  epsilon: 0.8110941400
04101  steps: 123  lr: 0.0001  evaluation reward: 2.46
episode: 1001 score: 0.0  memory length: 195529  epsilon: 0.8108506000
041062 steps: 123  lr: 0.0001  evaluation reward: 2.43
episode: 1002 score: 3.0  memory length: 195739  epsilon: 0.8104348000
```

```
041153  steps: 210  lr: 0.0001  evaluation reward: 2.44
episode: 1003  score: 3.0  memory length: 195987  epsilon: 0.8099437600
041259  steps: 248  lr: 0.0001  evaluation reward: 2.45
episode: 1004  score: 4.0  memory length: 196245  epsilon: 0.8094329200
04137   steps: 258  lr: 0.0001  evaluation reward: 2.47
episode: 1005  score: 3.0  memory length: 196472  epsilon: 0.8089834600
041468  steps: 227  lr: 0.0001  evaluation reward: 2.47
episode: 1006  score: 2.0  memory length: 196670  epsilon: 0.8085914200
041553  steps: 198  lr: 0.0001  evaluation reward: 2.46
episode: 1007  score: 2.0  memory length: 196886  epsilon: 0.8081637400
041646  steps: 216  lr: 0.0001  evaluation reward: 2.43
episode: 1008  score: 2.0  memory length: 197084  epsilon: 0.8077717000
041731  steps: 198  lr: 0.0001  evaluation reward: 2.4
episode: 1009  score: 2.0  memory length: 197264  epsilon: 0.8074153000
041808  steps: 180  lr: 0.0001  evaluation reward: 2.38
episode: 1010  score: 1.0  memory length: 197433  epsilon: 0.8070806800
041881  steps: 169  lr: 0.0001  evaluation reward: 2.35
episode: 1011  score: 1.0  memory length: 197602  epsilon: 0.8067460600
041954  steps: 169  lr: 0.0001  evaluation reward: 2.33
episode: 1012  score: 3.0  memory length: 197848  epsilon: 0.8062589800
042059  steps: 246  lr: 0.0001  evaluation reward: 2.33
episode: 1013  score: 2.0  memory length: 198048  epsilon: 0.8058629800
042145  steps: 200  lr: 0.0001  evaluation reward: 2.33
episode: 1014  score: 1.0  memory length: 198199  epsilon: 0.8055640000
04221  steps: 151  lr: 0.0001  evaluation reward: 2.34
episode: 1015  score: 6.0  memory length: 198574  epsilon: 0.8048215000
042371  steps: 375  lr: 0.0001  evaluation reward: 2.36
episode: 1016  score: 3.0  memory length: 198791  epsilon: 0.8043918400
042465  steps: 217  lr: 0.0001  evaluation reward: 2.35
episode: 1017  score: 5.0  memory length: 199107  epsilon: 0.8037661600
0426   steps: 316  lr: 0.0001  evaluation reward: 2.36
episode: 1018  score: 3.0  memory length: 199353  epsilon: 0.8032790800
042706  steps: 246  lr: 0.0001  evaluation reward: 2.36
episode: 1019  score: 2.0  memory length: 199569  epsilon: 0.8028514000
042799  steps: 216  lr: 0.0001  evaluation reward: 2.35
episode: 1020  score: 3.0  memory length: 199776  epsilon: 0.8024415400
042888  steps: 207  lr: 0.0001  evaluation reward: 2.36
episode: 1021  score: 3.0  memory length: 200043  epsilon: 0.8019128800
043003  steps: 267  lr: 4e-05  evaluation reward: 2.39
episode: 1022  score: 8.0  memory length: 200451  epsilon: 0.8011050400
043178  steps: 408  lr: 4e-05  evaluation reward: 2.44
episode: 1023  score: 1.0  memory length: 200620  epsilon: 0.8007704200
043251  steps: 169  lr: 4e-05  evaluation reward: 2.43
episode: 1024  score: 3.0  memory length: 200845  epsilon: 0.8003249200
043347  steps: 225  lr: 4e-05  evaluation reward: 2.45
episode: 1025  score: 0.0  memory length: 200968  epsilon: 0.8000813800
0434   steps: 123  lr: 4e-05  evaluation reward: 2.41
episode: 1026  score: 9.0  memory length: 201417  epsilon: 0.7991923600
043593  steps: 449  lr: 4e-05  evaluation reward: 2.47
episode: 1027  score: 6.0  memory length: 201754  epsilon: 0.7985251000
043738  steps: 337  lr: 4e-05  evaluation reward: 2.52
episode: 1028  score: 0.0  memory length: 201877  epsilon: 0.7982815600
043791  steps: 123  lr: 4e-05  evaluation reward: 2.49
```

```
episode: 1029  score: 3.0  memory length: 202103  epsilon: 0.7978340800
043888  steps: 226  lr: 4e-05  evaluation reward: 2.51
episode: 1030  score: 1.0  memory length: 202273  epsilon: 0.7974974800
043961  steps: 170  lr: 4e-05  evaluation reward: 2.51
episode: 1031  score: 4.0  memory length: 202552  epsilon: 0.7969450600
044081  steps: 279  lr: 4e-05  evaluation reward: 2.55
episode: 1032  score: 1.0  memory length: 202702  epsilon: 0.7966480600
044146  steps: 150  lr: 4e-05  evaluation reward: 2.56
episode: 1033  score: 6.0  memory length: 203096  epsilon: 0.7958679400
044315  steps: 394  lr: 4e-05  evaluation reward: 2.56
episode: 1034  score: 2.0  memory length: 203296  epsilon: 0.7954719400
044401  steps: 200  lr: 4e-05  evaluation reward: 2.56
episode: 1035  score: 3.0  memory length: 203562  epsilon: 0.7949452600
044515  steps: 266  lr: 4e-05  evaluation reward: 2.57
episode: 1036  score: 5.0  memory length: 203885  epsilon: 0.7943057200
044654  steps: 323  lr: 4e-05  evaluation reward: 2.59
episode: 1037  score: 2.0  memory length: 204085  epsilon: 0.7939097200
04474  steps: 200  lr: 4e-05  evaluation reward: 2.58
episode: 1038  score: 1.0  memory length: 204257  epsilon: 0.7935691600
044814  steps: 172  lr: 4e-05  evaluation reward: 2.55
episode: 1039  score: 6.0  memory length: 204646  epsilon: 0.7927989400
044981  steps: 389  lr: 4e-05  evaluation reward: 2.59
episode: 1040  score: 8.0  memory length: 205065  epsilon: 0.7919693200
045161  steps: 419  lr: 4e-05  evaluation reward: 2.65
episode: 1041  score: 4.0  memory length: 205360  epsilon: 0.7913852200
045288  steps: 295  lr: 4e-05  evaluation reward: 2.66
episode: 1042  score: 3.0  memory length: 205590  epsilon: 0.7909298200
045387  steps: 230  lr: 4e-05  evaluation reward: 2.66
episode: 1043  score: 3.0  memory length: 205834  epsilon: 0.7904467000
045492  steps: 244  lr: 4e-05  evaluation reward: 2.67
episode: 1044  score: 2.0  memory length: 206031  epsilon: 0.7900566400
045577  steps: 197  lr: 4e-05  evaluation reward: 2.67
episode: 1045  score: 4.0  memory length: 206307  epsilon: 0.7895101600
045695  steps: 276  lr: 4e-05  evaluation reward: 2.69
episode: 1046  score: 5.0  memory length: 206618  epsilon: 0.7888943800
045829  steps: 311  lr: 4e-05  evaluation reward: 2.7
episode: 1047  score: 1.0  memory length: 206788  epsilon: 0.7885577800
045902  steps: 170  lr: 4e-05  evaluation reward: 2.67
episode: 1048  score: 2.0  memory length: 206968  epsilon: 0.7882013800
045979  steps: 180  lr: 4e-05  evaluation reward: 2.66
episode: 1049  score: 5.0  memory length: 207272  epsilon: 0.7875994600
04611  steps: 304  lr: 4e-05  evaluation reward: 2.69
episode: 1050  score: 3.0  memory length: 207501  epsilon: 0.7871460400
046209  steps: 229  lr: 4e-05  evaluation reward: 2.7
episode: 1051  score: 2.0  memory length: 207701  epsilon: 0.7867500400
046294  steps: 200  lr: 4e-05  evaluation reward: 2.67
episode: 1052  score: 3.0  memory length: 207911  epsilon: 0.7863342400
046385  steps: 210  lr: 4e-05  evaluation reward: 2.7
episode: 1053  score: 3.0  memory length: 208137  epsilon: 0.7858867600
046482  steps: 226  lr: 4e-05  evaluation reward: 2.73
episode: 1054  score: 1.0  memory length: 208288  epsilon: 0.7855877800
046547  steps: 151  lr: 4e-05  evaluation reward: 2.7
episode: 1055  score: 2.0  memory length: 208487  epsilon: 0.7851937600
```

```
046632  steps: 199  lr: 4e-05  evaluation reward: 2.67
episode: 1056  score: 6.0  memory length: 208836  epsilon: 0.7845027400
046782  steps: 349  lr: 4e-05  evaluation reward: 2.68
episode: 1057  score: 5.0  memory length: 209155  epsilon: 0.7838711200
04692   steps: 319  lr: 4e-05  evaluation reward: 2.73
episode: 1058  score: 3.0  memory length: 209423  epsilon: 0.7833404800
047035  steps: 268  lr: 4e-05  evaluation reward: 2.74
episode: 1059  score: 3.0  memory length: 209653  epsilon: 0.7828850800
047134  steps: 230  lr: 4e-05  evaluation reward: 2.73
episode: 1060  score: 0.0  memory length: 209775  epsilon: 0.7826435200
047186  steps: 122  lr: 4e-05  evaluation reward: 2.73
episode: 1061  score: 6.0  memory length: 210169  epsilon: 0.7818634000
047355  steps: 394  lr: 4e-05  evaluation reward: 2.79
episode: 1062  score: 3.0  memory length: 210413  epsilon: 0.7813802800
04746   steps: 244  lr: 4e-05  evaluation reward: 2.79
episode: 1063  score: 5.0  memory length: 210739  epsilon: 0.7807348000
0476    steps: 326  lr: 4e-05  evaluation reward: 2.79
episode: 1064  score: 6.0  memory length: 211132  epsilon: 0.7799566600
047769  steps: 393  lr: 4e-05  evaluation reward: 2.82
episode: 1065  score: 3.0  memory length: 211377  epsilon: 0.7794715600
047875  steps: 245  lr: 4e-05  evaluation reward: 2.85
episode: 1066  score: 3.0  memory length: 211602  epsilon: 0.7790260600
047971  steps: 225  lr: 4e-05  evaluation reward: 2.87
episode: 1067  score: 4.0  memory length: 211881  epsilon: 0.7784736400
048091  steps: 279  lr: 4e-05  evaluation reward: 2.87
episode: 1068  score: 3.0  memory length: 212144  epsilon: 0.7779529000
048204  steps: 263  lr: 4e-05  evaluation reward: 2.89
episode: 1069  score: 2.0  memory length: 212342  epsilon: 0.7775608600
048289  steps: 198  lr: 4e-05  evaluation reward: 2.88
episode: 1070  score: 5.0  memory length: 212671  epsilon: 0.7769094400
048431  steps: 329  lr: 4e-05  evaluation reward: 2.91
episode: 1071  score: 2.0  memory length: 212850  epsilon: 0.7765550200
048508  steps: 179  lr: 4e-05  evaluation reward: 2.9
episode: 1072  score: 5.0  memory length: 213138  epsilon: 0.7759847800
048632  steps: 288  lr: 4e-05  evaluation reward: 2.91
episode: 1073  score: 4.0  memory length: 213416  epsilon: 0.7754343400
048751  steps: 278  lr: 4e-05  evaluation reward: 2.91
episode: 1074  score: 5.0  memory length: 213698  epsilon: 0.7748759800
048872  steps: 282  lr: 4e-05  evaluation reward: 2.94
episode: 1075  score: 3.0  memory length: 213927  epsilon: 0.7744225600
048971  steps: 229  lr: 4e-05  evaluation reward: 2.94
episode: 1076  score: 2.0  memory length: 214127  epsilon: 0.7740265600
049057  steps: 200  lr: 4e-05  evaluation reward: 2.95
episode: 1077  score: 4.0  memory length: 214420  epsilon: 0.7734464200
049183  steps: 293  lr: 4e-05  evaluation reward: 2.99
episode: 1078  score: 0.0  memory length: 214543  epsilon: 0.7732028800
049235  steps: 123  lr: 4e-05  evaluation reward: 2.98
episode: 1079  score: 3.0  memory length: 214753  epsilon: 0.7727870800
049326  steps: 210  lr: 4e-05  evaluation reward: 2.99
episode: 1080  score: 4.0  memory length: 215028  epsilon: 0.7722425800
049444  steps: 275  lr: 4e-05  evaluation reward: 3.0
episode: 1081  score: 2.0  memory length: 215226  epsilon: 0.7718505400
049529  steps: 198  lr: 4e-05  evaluation reward: 3.02
```

```
episode: 1082  score: 7.0  memory length: 215649  epsilon: 0.7710130000
049711  steps: 423  lr: 4e-05  evaluation reward: 3.07
episode: 1083  score: 4.0  memory length: 215890  epsilon: 0.7705358200
049814  steps: 241  lr: 4e-05  evaluation reward: 3.08
episode: 1084  score: 0.0  memory length: 216013  epsilon: 0.7702922800
049867  steps: 123  lr: 4e-05  evaluation reward: 3.07
episode: 1085  score: 1.0  memory length: 216184  epsilon: 0.7699537000
049941  steps: 171  lr: 4e-05  evaluation reward: 3.08
episode: 1086  score: 5.0  memory length: 216522  epsilon: 0.7692844600
050086  steps: 338  lr: 4e-05  evaluation reward: 3.09
episode: 1087  score: 2.0  memory length: 216720  epsilon: 0.7688924200
050171  steps: 198  lr: 4e-05  evaluation reward: 3.09
episode: 1088  score: 1.0  memory length: 216871  epsilon: 0.7685934400
050236  steps: 151  lr: 4e-05  evaluation reward: 3.09
episode: 1089  score: 8.0  memory length: 217311  epsilon: 0.7677222400
050425  steps: 440  lr: 4e-05  evaluation reward: 3.15
episode: 1090  score: 4.0  memory length: 217588  epsilon: 0.7671737800
050544  steps: 277  lr: 4e-05  evaluation reward: 3.19
episode: 1091  score: 2.0  memory length: 217805  epsilon: 0.7667441200
050638  steps: 217  lr: 4e-05  evaluation reward: 3.15
episode: 1092  score: 4.0  memory length: 218097  epsilon: 0.7661659600
050763  steps: 292  lr: 4e-05  evaluation reward: 3.17
episode: 1093  score: 7.0  memory length: 218520  epsilon: 0.7653284200
050945  steps: 423  lr: 4e-05  evaluation reward: 3.22
episode: 1094  score: 2.0  memory length: 218718  epsilon: 0.7649363800
05103  steps: 198  lr: 4e-05  evaluation reward: 3.21
episode: 1095  score: 6.0  memory length: 219102  epsilon: 0.7641760600
051195  steps: 384  lr: 4e-05  evaluation reward: 3.24
episode: 1096  score: 0.0  memory length: 219224  epsilon: 0.7639345000
051248  steps: 122  lr: 4e-05  evaluation reward: 3.2
episode: 1097  score: 4.0  memory length: 219523  epsilon: 0.7633424800
051376  steps: 299  lr: 4e-05  evaluation reward: 3.21
episode: 1098  score: 4.0  memory length: 219803  epsilon: 0.7627880800
051496  steps: 280  lr: 4e-05  evaluation reward: 3.2
episode: 1099  score: 4.0  memory length: 220116  epsilon: 0.7621683400
051631  steps: 313  lr: 4e-05  evaluation reward: 3.22
episode: 1100  score: 3.0  memory length: 220342  epsilon: 0.7617208600
051728  steps: 226  lr: 4e-05  evaluation reward: 3.25
episode: 1101  score: 3.0  memory length: 220590  epsilon: 0.7612298200
051835  steps: 248  lr: 4e-05  evaluation reward: 3.28
episode: 1102  score: 1.0  memory length: 220759  epsilon: 0.7608952000
051907  steps: 169  lr: 4e-05  evaluation reward: 3.26
episode: 1103  score: 6.0  memory length: 221139  epsilon: 0.7601428000
052071  steps: 380  lr: 4e-05  evaluation reward: 3.29
episode: 1104  score: 2.0  memory length: 221356  epsilon: 0.7597131400
052164  steps: 217  lr: 4e-05  evaluation reward: 3.27
episode: 1105  score: 3.0  memory length: 221603  epsilon: 0.7592240800
05227  steps: 247  lr: 4e-05  evaluation reward: 3.27
episode: 1106  score: 3.0  memory length: 221834  epsilon: 0.7587667000
052369  steps: 231  lr: 4e-05  evaluation reward: 3.28
episode: 1107  score: 4.0  memory length: 222130  epsilon: 0.7581806200
052497  steps: 296  lr: 4e-05  evaluation reward: 3.3
episode: 1108  score: 1.0  memory length: 222301  epsilon: 0.7578420400
```

```
05257    steps: 171    lr: 4e-05    evaluation reward: 3.29
episode: 1109    score: 1.0    memory length: 222470    epsilon: 0.7575074200
052643    steps: 169    lr: 4e-05    evaluation reward: 3.28
episode: 1110    score: 1.0    memory length: 222639    epsilon: 0.7571728000
052715    steps: 169    lr: 4e-05    evaluation reward: 3.28
episode: 1111    score: 6.0    memory length: 222992    epsilon: 0.7564738600
052867    steps: 353    lr: 4e-05    evaluation reward: 3.33
episode: 1112    score: 0.0    memory length: 223114    epsilon: 0.7562323000
05292    steps: 122    lr: 4e-05    evaluation reward: 3.3
episode: 1113    score: 2.0    memory length: 223313    epsilon: 0.7558382800
053005    steps: 199    lr: 4e-05    evaluation reward: 3.3
episode: 1114    score: 5.0    memory length: 223654    epsilon: 0.7551631000
053152    steps: 341    lr: 4e-05    evaluation reward: 3.34
episode: 1115    score: 6.0    memory length: 224011    epsilon: 0.7544562400
053305    steps: 357    lr: 4e-05    evaluation reward: 3.34
episode: 1116    score: 2.0    memory length: 224192    epsilon: 0.7540978600
053383    steps: 181    lr: 4e-05    evaluation reward: 3.33
episode: 1117    score: 3.0    memory length: 224456    epsilon: 0.7535751400
053496    steps: 264    lr: 4e-05    evaluation reward: 3.31
episode: 1118    score: 3.0    memory length: 224702    epsilon: 0.7530880600
053602    steps: 246    lr: 4e-05    evaluation reward: 3.31
episode: 1119    score: 3.0    memory length: 224949    epsilon: 0.7525990000
053708    steps: 247    lr: 4e-05    evaluation reward: 3.32
episode: 1120    score: 1.0    memory length: 225099    epsilon: 0.7523020000
053773    steps: 150    lr: 4e-05    evaluation reward: 3.3
episode: 1121    score: 5.0    memory length: 225440    epsilon: 0.7516268200
053919    steps: 341    lr: 4e-05    evaluation reward: 3.32
episode: 1122    score: 4.0    memory length: 225714    epsilon: 0.7510843000
054037    steps: 274    lr: 4e-05    evaluation reward: 3.28
episode: 1123    score: 3.0    memory length: 225941    epsilon: 0.7506348400
054135    steps: 227    lr: 4e-05    evaluation reward: 3.3
episode: 1124    score: 8.0    memory length: 226386    epsilon: 0.7497537400
054326    steps: 445    lr: 4e-05    evaluation reward: 3.35
episode: 1125    score: 6.0    memory length: 226699    epsilon: 0.7491340000
05446    steps: 313    lr: 4e-05    evaluation reward: 3.41
episode: 1126    score: 5.0    memory length: 227004    epsilon: 0.7485301000
054592    steps: 305    lr: 4e-05    evaluation reward: 3.37
episode: 1127    score: 2.0    memory length: 227221    epsilon: 0.7481004400
054685    steps: 217    lr: 4e-05    evaluation reward: 3.33
episode: 1128    score: 2.0    memory length: 227440    epsilon: 0.7476668200
054779    steps: 219    lr: 4e-05    evaluation reward: 3.35
episode: 1129    score: 1.0    memory length: 227608    epsilon: 0.7473341800
054851    steps: 168    lr: 4e-05    evaluation reward: 3.33
episode: 1130    score: 9.0    memory length: 228065    epsilon: 0.7464293200
055048    steps: 457    lr: 4e-05    evaluation reward: 3.41
episode: 1131    score: 2.0    memory length: 228264    epsilon: 0.7460353000
055133    steps: 199    lr: 4e-05    evaluation reward: 3.39
episode: 1132    score: 3.0    memory length: 228490    epsilon: 0.7455878200
05523    steps: 226    lr: 4e-05    evaluation reward: 3.41
episode: 1133    score: 3.0    memory length: 228718    epsilon: 0.7451363800
055328    steps: 228    lr: 4e-05    evaluation reward: 3.38
episode: 1134    score: 2.0    memory length: 228935    epsilon: 0.7447067200
055422    steps: 217    lr: 4e-05    evaluation reward: 3.38
```



```
episode: 1135  score: 4.0  memory length: 229252  epsilon: 0.7440790600
055558  steps: 317  lr: 4e-05  evaluation reward: 3.39
episode: 1136  score: 5.0  memory length: 229560  epsilon: 0.7434692200
055569  steps: 308  lr: 4e-05  evaluation reward: 3.39
episode: 1137  score: 2.0  memory length: 229781  epsilon: 0.7430316400
055785  steps: 221  lr: 4e-05  evaluation reward: 3.39
episode: 1138  score: 2.0  memory length: 230000  epsilon: 0.7425980200
055879  steps: 219  lr: 4e-05  evaluation reward: 3.4
episode: 1139  score: 7.0  memory length: 230385  epsilon: 0.7418357200
056045  steps: 385  lr: 4e-05  evaluation reward: 3.41
episode: 1140  score: 4.0  memory length: 230682  epsilon: 0.7412476600
056173  steps: 297  lr: 4e-05  evaluation reward: 3.37
episode: 1141  score: 5.0  memory length: 230988  epsilon: 0.7406417800
056304  steps: 306  lr: 4e-05  evaluation reward: 3.38
episode: 1142  score: 7.0  memory length: 231398  epsilon: 0.7398299800
05648  steps: 410  lr: 4e-05  evaluation reward: 3.42
episode: 1143  score: 7.0  memory length: 231825  epsilon: 0.7389845200
056664  steps: 427  lr: 4e-05  evaluation reward: 3.46
episode: 1144  score: 2.0  memory length: 232025  epsilon: 0.7385885200
05675  steps: 200  lr: 4e-05  evaluation reward: 3.46
episode: 1145  score: 1.0  memory length: 232197  epsilon: 0.7382479600
056824  steps: 172  lr: 4e-05  evaluation reward: 3.43
episode: 1146  score: 2.0  memory length: 232399  epsilon: 0.7378480000
056911  steps: 202  lr: 4e-05  evaluation reward: 3.4
episode: 1147  score: 5.0  memory length: 232725  epsilon: 0.7372025200
057051  steps: 326  lr: 4e-05  evaluation reward: 3.44
episode: 1148  score: 0.0  memory length: 232848  epsilon: 0.7369589800
057104  steps: 123  lr: 4e-05  evaluation reward: 3.42
episode: 1149  score: 1.0  memory length: 232998  epsilon: 0.7366619800
057168  steps: 150  lr: 4e-05  evaluation reward: 3.38
episode: 1150  score: 2.0  memory length: 233178  epsilon: 0.7363055800
057245  steps: 180  lr: 4e-05  evaluation reward: 3.37
episode: 1151  score: 6.0  memory length: 233552  epsilon: 0.7355650600
057406  steps: 374  lr: 4e-05  evaluation reward: 3.41
episode: 1152  score: 8.0  memory length: 233990  epsilon: 0.7346978200
057594  steps: 438  lr: 4e-05  evaluation reward: 3.46
episode: 1153  score: 4.0  memory length: 234307  epsilon: 0.7340701600
057731  steps: 317  lr: 4e-05  evaluation reward: 3.47
episode: 1154  score: 5.0  memory length: 234617  epsilon: 0.7334563600
057864  steps: 310  lr: 4e-05  evaluation reward: 3.51
episode: 1155  score: 7.0  memory length: 234980  epsilon: 0.7327376200
05802  steps: 363  lr: 4e-05  evaluation reward: 3.56
episode: 1156  score: 2.0  memory length: 235178  epsilon: 0.7323455800
058105  steps: 198  lr: 4e-05  evaluation reward: 3.52
episode: 1157  score: 1.0  memory length: 235328  epsilon: 0.7320485800
05817  steps: 150  lr: 4e-05  evaluation reward: 3.48
episode: 1158  score: 4.0  memory length: 235589  epsilon: 0.7315318000
058282  steps: 261  lr: 4e-05  evaluation reward: 3.49
episode: 1159  score: 5.0  memory length: 235900  epsilon: 0.7309160200
058415  steps: 311  lr: 4e-05  evaluation reward: 3.51
episode: 1160  score: 0.0  memory length: 236022  epsilon: 0.7306744600
058468  steps: 122  lr: 4e-05  evaluation reward: 3.51
episode: 1161  score: 7.0  memory length: 236411  epsilon: 0.7299042400
```

```
058635    steps: 389    lr: 4e-05    evaluation reward: 3.52
episode: 1162    score: 4.0    memory length: 236707    epsilon: 0.7293181600
058762    steps: 296    lr: 4e-05    evaluation reward: 3.53
episode: 1163    score: 8.0    memory length: 237034    epsilon: 0.7286707000
058903    steps: 327    lr: 4e-05    evaluation reward: 3.56
episode: 1164    score: 6.0    memory length: 237363    epsilon: 0.7280192800
059044    steps: 329    lr: 4e-05    evaluation reward: 3.56
episode: 1165    score: 6.0    memory length: 237691    epsilon: 0.7273698400
059185    steps: 328    lr: 4e-05    evaluation reward: 3.59
episode: 1166    score: 0.0    memory length: 237814    epsilon: 0.7271263000
059238    steps: 123    lr: 4e-05    evaluation reward: 3.56
episode: 1167    score: 3.0    memory length: 238040    epsilon: 0.7266788200
059335    steps: 226    lr: 4e-05    evaluation reward: 3.55
episode: 1168    score: 1.0    memory length: 238209    epsilon: 0.7263442000
059408    steps: 169    lr: 4e-05    evaluation reward: 3.53
episode: 1169    score: 2.0    memory length: 238390    epsilon: 0.7259858200
059486    steps: 181    lr: 4e-05    evaluation reward: 3.53
episode: 1170    score: 2.0    memory length: 238588    epsilon: 0.7255937800
059571    steps: 198    lr: 4e-05    evaluation reward: 3.5
episode: 1171    score: 8.0    memory length: 238998    epsilon: 0.7247819800
059747    steps: 410    lr: 4e-05    evaluation reward: 3.56
episode: 1172    score: 4.0    memory length: 239273    epsilon: 0.7242374800
059865    steps: 275    lr: 4e-05    evaluation reward: 3.55
episode: 1173    score: 7.0    memory length: 239679    epsilon: 0.7234336000
06004    steps: 406    lr: 4e-05    evaluation reward: 3.58
episode: 1174    score: 4.0    memory length: 239974    epsilon: 0.7228495000
060167    steps: 295    lr: 4e-05    evaluation reward: 3.57
episode: 1175    score: 4.0    memory length: 240248    epsilon: 0.7223069800
060284    steps: 274    lr: 4e-05    evaluation reward: 3.58
episode: 1176    score: 3.0    memory length: 240473    epsilon: 0.7218614800
060381    steps: 225    lr: 4e-05    evaluation reward: 3.59
episode: 1177    score: 1.0    memory length: 240623    epsilon: 0.7215644800
060446    steps: 150    lr: 4e-05    evaluation reward: 3.56
episode: 1178    score: 3.0    memory length: 240868    epsilon: 0.7210793800
060551    steps: 245    lr: 4e-05    evaluation reward: 3.59
episode: 1179    score: 3.0    memory length: 241094    epsilon: 0.7206319000
060648    steps: 226    lr: 4e-05    evaluation reward: 3.59
episode: 1180    score: 3.0    memory length: 241307    epsilon: 0.7202101600
06074    steps: 213    lr: 4e-05    evaluation reward: 3.58
episode: 1181    score: 3.0    memory length: 241554    epsilon: 0.7197211000
060846    steps: 247    lr: 4e-05    evaluation reward: 3.59
episode: 1182    score: 5.0    memory length: 241919    epsilon: 0.7189984000
061003    steps: 365    lr: 4e-05    evaluation reward: 3.57
episode: 1183    score: 4.0    memory length: 242212    epsilon: 0.7184182600
061129    steps: 293    lr: 4e-05    evaluation reward: 3.57
episode: 1184    score: 1.0    memory length: 242362    epsilon: 0.7181212600
061193    steps: 150    lr: 4e-05    evaluation reward: 3.58
episode: 1185    score: 4.0    memory length: 242634    epsilon: 0.7175827000
06131    steps: 272    lr: 4e-05    evaluation reward: 3.61
episode: 1186    score: 3.0    memory length: 242861    epsilon: 0.7171332400
061408    steps: 227    lr: 4e-05    evaluation reward: 3.59
episode: 1187    score: 4.0    memory length: 243124    epsilon: 0.7166125000
061521    steps: 263    lr: 4e-05    evaluation reward: 3.61
```

```
episode: 1188  score: 6.0  memory length: 243497  epsilon: 0.7158739600
061681  steps: 373  lr: 4e-05  evaluation reward: 3.66
episode: 1189  score: 2.0  memory length: 243694  epsilon: 0.7154839000
061766  steps: 197  lr: 4e-05  evaluation reward: 3.6
episode: 1190  score: 9.0  memory length: 244174  epsilon: 0.7145335000
061972  steps: 480  lr: 4e-05  evaluation reward: 3.65
episode: 1191  score: 4.0  memory length: 244450  epsilon: 0.7139870200
062091  steps: 276  lr: 4e-05  evaluation reward: 3.67
episode: 1192  score: 5.0  memory length: 244757  epsilon: 0.7133791600
062223  steps: 307  lr: 4e-05  evaluation reward: 3.68
episode: 1193  score: 7.0  memory length: 245185  epsilon: 0.7125317200
062407  steps: 428  lr: 4e-05  evaluation reward: 3.68
episode: 1194  score: 3.0  memory length: 245412  epsilon: 0.7120822600
062504  steps: 227  lr: 4e-05  evaluation reward: 3.69
episode: 1195  score: 3.0  memory length: 245660  epsilon: 0.7115912200
062611  steps: 248  lr: 4e-05  evaluation reward: 3.66
episode: 1196  score: 2.0  memory length: 245840  epsilon: 0.7112348200
062688  steps: 180  lr: 4e-05  evaluation reward: 3.68
episode: 1197  score: 4.0  memory length: 246095  epsilon: 0.7107299200
062798  steps: 255  lr: 4e-05  evaluation reward: 3.68
episode: 1198  score: 4.0  memory length: 246369  epsilon: 0.7101874000
062915  steps: 274  lr: 4e-05  evaluation reward: 3.68
episode: 1199  score: 6.0  memory length: 246742  epsilon: 0.7094488600
063076  steps: 373  lr: 4e-05  evaluation reward: 3.7
episode: 1200  score: 7.0  memory length: 247119  epsilon: 0.7087024000
063238  steps: 377  lr: 4e-05  evaluation reward: 3.74
episode: 1201  score: 0.0  memory length: 247242  epsilon: 0.7084588600
063291  steps: 123  lr: 4e-05  evaluation reward: 3.71
episode: 1202  score: 8.0  memory length: 247678  epsilon: 0.7075955800
063478  steps: 436  lr: 4e-05  evaluation reward: 3.78
episode: 1203  score: 3.0  memory length: 247904  epsilon: 0.7071481000
063575  steps: 226  lr: 4e-05  evaluation reward: 3.75
episode: 1204  score: 3.0  memory length: 248148  epsilon: 0.7066649800
06368  steps: 244  lr: 4e-05  evaluation reward: 3.76
episode: 1205  score: 3.0  memory length: 248396  epsilon: 0.7061739400
063787  steps: 248  lr: 4e-05  evaluation reward: 3.76
episode: 1206  score: 2.0  memory length: 248613  epsilon: 0.7057442800
06388  steps: 217  lr: 4e-05  evaluation reward: 3.75
episode: 1207  score: 3.0  memory length: 248859  epsilon: 0.7052572000
063986  steps: 246  lr: 4e-05  evaluation reward: 3.74
episode: 1208  score: 5.0  memory length: 249205  epsilon: 0.7045721200
064134  steps: 346  lr: 4e-05  evaluation reward: 3.78
episode: 1209  score: 2.0  memory length: 249407  epsilon: 0.7041721600
064221  steps: 202  lr: 4e-05  evaluation reward: 3.79
episode: 1210  score: 2.0  memory length: 249628  epsilon: 0.7037345800
064316  steps: 221  lr: 4e-05  evaluation reward: 3.8
episode: 1211  score: 2.0  memory length: 249828  epsilon: 0.7033385800
064402  steps: 200  lr: 4e-05  evaluation reward: 3.76
episode: 1212  score: 3.0  memory length: 250055  epsilon: 0.7028891200
0645  steps: 227  lr: 4e-05  evaluation reward: 3.79
episode: 1213  score: 0.0  memory length: 250178  epsilon: 0.7026455800
064553  steps: 123  lr: 4e-05  evaluation reward: 3.77
episode: 1214  score: 2.0  memory length: 250380  epsilon: 0.7022456200
```

```
06464    steps: 202    lr: 4e-05    evaluation reward: 3.74
episode: 1215    score: 3.0    memory length: 250606    epsilon: 0.7017981400
064737    steps: 226    lr: 4e-05    evaluation reward: 3.71
episode: 1216    score: 3.0    memory length: 250833    epsilon: 0.7013486800
064834    steps: 227    lr: 4e-05    evaluation reward: 3.72
episode: 1217    score: 4.0    memory length: 251095    epsilon: 0.7008299200
064947    steps: 262    lr: 4e-05    evaluation reward: 3.73
episode: 1218    score: 1.0    memory length: 251246    epsilon: 0.7005309400
065012    steps: 151    lr: 4e-05    evaluation reward: 3.71
episode: 1219    score: 3.0    memory length: 251474    epsilon: 0.7000795000
06511    steps: 228    lr: 4e-05    evaluation reward: 3.71
episode: 1220    score: 10.0    memory length: 251965    epsilon: 0.699107320
0065321    steps: 491    lr: 4e-05    evaluation reward: 3.8
episode: 1221    score: 5.0    memory length: 252264    epsilon: 0.6985153000
065449    steps: 299    lr: 4e-05    evaluation reward: 3.8
episode: 1222    score: 4.0    memory length: 252524    epsilon: 0.6980005000
065561    steps: 260    lr: 4e-05    evaluation reward: 3.8
episode: 1223    score: 2.0    memory length: 252740    epsilon: 0.6975728200
065654    steps: 216    lr: 4e-05    evaluation reward: 3.79
episode: 1224    score: 4.0    memory length: 253008    epsilon: 0.6970421800
065769    steps: 268    lr: 4e-05    evaluation reward: 3.75
episode: 1225    score: 3.0    memory length: 253236    epsilon: 0.6965907400
065867    steps: 228    lr: 4e-05    evaluation reward: 3.72
episode: 1226    score: 2.0    memory length: 253434    epsilon: 0.6961987000
065952    steps: 198    lr: 4e-05    evaluation reward: 3.69
episode: 1227    score: 2.0    memory length: 253633    epsilon: 0.6958046800
066038    steps: 199    lr: 4e-05    evaluation reward: 3.69
episode: 1228    score: 10.0    memory length: 254039    epsilon: 0.695000800
0066212    steps: 406    lr: 4e-05    evaluation reward: 3.77
episode: 1229    score: 7.0    memory length: 254425    epsilon: 0.6942365200
066378    steps: 386    lr: 4e-05    evaluation reward: 3.83
episode: 1230    score: 5.0    memory length: 254767    epsilon: 0.6935593600
066525    steps: 342    lr: 4e-05    evaluation reward: 3.79
episode: 1231    score: 1.0    memory length: 254937    epsilon: 0.6932227600
066598    steps: 170    lr: 4e-05    evaluation reward: 3.78
episode: 1232    score: 4.0    memory length: 255215    epsilon: 0.6926723200
066718    steps: 278    lr: 4e-05    evaluation reward: 3.79
episode: 1233    score: 3.0    memory length: 255464    epsilon: 0.6921793000
066825    steps: 249    lr: 4e-05    evaluation reward: 3.79
episode: 1234    score: 5.0    memory length: 255785    epsilon: 0.6915437200
066963    steps: 321    lr: 4e-05    evaluation reward: 3.82
episode: 1235    score: 4.0    memory length: 256080    epsilon: 0.6909596200
06709    steps: 295    lr: 4e-05    evaluation reward: 3.82
episode: 1236    score: 6.0    memory length: 256494    epsilon: 0.6901399000
067268    steps: 414    lr: 4e-05    evaluation reward: 3.83
episode: 1237    score: 1.0    memory length: 256644    epsilon: 0.6898429000
067332    steps: 150    lr: 4e-05    evaluation reward: 3.82
episode: 1238    score: 2.0    memory length: 256842    epsilon: 0.6894508600
067417    steps: 198    lr: 4e-05    evaluation reward: 3.82
episode: 1239    score: 1.0    memory length: 256992    epsilon: 0.6891538600
067482    steps: 150    lr: 4e-05    evaluation reward: 3.76
episode: 1240    score: 0.0    memory length: 257114    epsilon: 0.6889123000
067534    steps: 122    lr: 4e-05    evaluation reward: 3.72
```

```
episode: 1241  score: 7.0  memory length: 257528  epsilon: 0.6880925800
067712  steps: 414  lr: 4e-05  evaluation reward: 3.74
episode: 1242  score: 5.0  memory length: 257826  epsilon: 0.6875025400
06784  steps: 298  lr: 4e-05  evaluation reward: 3.72
episode: 1243  score: 8.0  memory length: 258121  epsilon: 0.6869184400
067967  steps: 295  lr: 4e-05  evaluation reward: 3.73
episode: 1244  score: 3.0  memory length: 258334  epsilon: 0.6864967000
068058  steps: 213  lr: 4e-05  evaluation reward: 3.74
episode: 1245  score: 4.0  memory length: 258599  epsilon: 0.6859720000
068172  steps: 265  lr: 4e-05  evaluation reward: 3.77
episode: 1246  score: 4.0  memory length: 258902  epsilon: 0.6853720600
068303  steps: 303  lr: 4e-05  evaluation reward: 3.79
episode: 1247  score: 3.0  memory length: 259167  epsilon: 0.6848473600
068417  steps: 265  lr: 4e-05  evaluation reward: 3.77
episode: 1248  score: 7.0  memory length: 259571  epsilon: 0.6840474400
06859  steps: 404  lr: 4e-05  evaluation reward: 3.84
episode: 1249  score: 4.0  memory length: 259833  epsilon: 0.6835286800
068703  steps: 262  lr: 4e-05  evaluation reward: 3.87
episode: 1250  score: 2.0  memory length: 260031  epsilon: 0.6831366400
068788  steps: 198  lr: 4e-05  evaluation reward: 3.87
episode: 1251  score: 6.0  memory length: 260396  epsilon: 0.6824139400
068945  steps: 365  lr: 4e-05  evaluation reward: 3.87
episode: 1252  score: 4.0  memory length: 260674  epsilon: 0.6818635000
069064  steps: 278  lr: 4e-05  evaluation reward: 3.83
episode: 1253  score: 7.0  memory length: 261073  epsilon: 0.6810734800
069236  steps: 399  lr: 4e-05  evaluation reward: 3.86
episode: 1254  score: 3.0  memory length: 261320  epsilon: 0.6805844200
069342  steps: 247  lr: 4e-05  evaluation reward: 3.84
episode: 1255  score: 7.0  memory length: 261724  epsilon: 0.6797845000
069516  steps: 404  lr: 4e-05  evaluation reward: 3.84
episode: 1256  score: 7.0  memory length: 262099  epsilon: 0.6790420000
069677  steps: 375  lr: 4e-05  evaluation reward: 3.89
episode: 1257  score: 3.0  memory length: 262309  epsilon: 0.6786262000
069767  steps: 210  lr: 4e-05  evaluation reward: 3.91
episode: 1258  score: 3.0  memory length: 262557  epsilon: 0.6781351600
069874  steps: 248  lr: 4e-05  evaluation reward: 3.9
episode: 1259  score: 9.0  memory length: 263028  epsilon: 0.6772025800
070076  steps: 471  lr: 4e-05  evaluation reward: 3.94
episode: 1260  score: 5.0  memory length: 263348  epsilon: 0.6765689800
070214  steps: 320  lr: 4e-05  evaluation reward: 3.99
episode: 1261  score: 3.0  memory length: 263576  epsilon: 0.6761175400
070312  steps: 228  lr: 4e-05  evaluation reward: 3.95
episode: 1262  score: 2.0  memory length: 263776  epsilon: 0.6757215400
070398  steps: 200  lr: 4e-05  evaluation reward: 3.93
episode: 1263  score: 4.0  memory length: 264054  epsilon: 0.6751711000
070517  steps: 278  lr: 4e-05  evaluation reward: 3.89
episode: 1264  score: 4.0  memory length: 264328  epsilon: 0.6746285800
070635  steps: 274  lr: 4e-05  evaluation reward: 3.87
episode: 1265  score: 3.0  memory length: 264593  epsilon: 0.6741038800
070749  steps: 265  lr: 4e-05  evaluation reward: 3.84
episode: 1266  score: 5.0  memory length: 264897  epsilon: 0.6735019600
07088  steps: 304  lr: 4e-05  evaluation reward: 3.89
episode: 1267  score: 4.0  memory length: 265141  epsilon: 0.6730188400
```

```
070984 steps: 244 lr: 4e-05 evaluation reward: 3.9
episode: 1268 score: 2.0 memory length: 265339 epsilon: 0.6726268000
07107 steps: 198 lr: 4e-05 evaluation reward: 3.91
episode: 1269 score: 3.0 memory length: 265551 epsilon: 0.6722070400
071161 steps: 212 lr: 4e-05 evaluation reward: 3.92
episode: 1270 score: 5.0 memory length: 265876 epsilon: 0.6715635400
0713 steps: 325 lr: 4e-05 evaluation reward: 3.95
episode: 1271 score: 6.0 memory length: 266215 epsilon: 0.6708923200
071446 steps: 339 lr: 4e-05 evaluation reward: 3.93
episode: 1272 score: 12.0 memory length: 266745 epsilon: 0.669842920
0071674 steps: 530 lr: 4e-05 evaluation reward: 4.01
episode: 1273 score: 2.0 memory length: 266929 epsilon: 0.6694786000
071753 steps: 184 lr: 4e-05 evaluation reward: 3.96
episode: 1274 score: 1.0 memory length: 267079 epsilon: 0.6691816000
071817 steps: 150 lr: 4e-05 evaluation reward: 3.93
episode: 1275 score: 1.0 memory length: 267230 epsilon: 0.6688826200
071882 steps: 151 lr: 4e-05 evaluation reward: 3.9
episode: 1276 score: 3.0 memory length: 267458 epsilon: 0.6684311800
07198 steps: 228 lr: 4e-05 evaluation reward: 3.9
episode: 1277 score: 4.0 memory length: 267758 epsilon: 0.6678371800
072109 steps: 300 lr: 4e-05 evaluation reward: 3.93
episode: 1278 score: 6.0 memory length: 268108 epsilon: 0.6671441800
07226 steps: 350 lr: 4e-05 evaluation reward: 3.96
episode: 1279 score: 3.0 memory length: 268359 epsilon: 0.6666472000
072368 steps: 251 lr: 4e-05 evaluation reward: 3.96
episode: 1280 score: 3.0 memory length: 268606 epsilon: 0.6661581400
072474 steps: 247 lr: 4e-05 evaluation reward: 3.96
episode: 1281 score: 1.0 memory length: 268775 epsilon: 0.6658235200
072546 steps: 169 lr: 4e-05 evaluation reward: 3.94
episode: 1282 score: 4.0 memory length: 269050 epsilon: 0.6652790200
072665 steps: 275 lr: 4e-05 evaluation reward: 3.93
episode: 1283 score: 6.0 memory length: 269423 epsilon: 0.6645404800
072825 steps: 373 lr: 4e-05 evaluation reward: 3.95
episode: 1284 score: 4.0 memory length: 269680 epsilon: 0.6640316200
072935 steps: 257 lr: 4e-05 evaluation reward: 3.98
episode: 1285 score: 4.0 memory length: 269934 epsilon: 0.6635287000
073045 steps: 254 lr: 4e-05 evaluation reward: 3.98
episode: 1286 score: 5.0 memory length: 270245 epsilon: 0.6629129200
073178 steps: 311 lr: 4e-05 evaluation reward: 4.0
episode: 1287 score: 3.0 memory length: 270495 epsilon: 0.6624179200
073286 steps: 250 lr: 4e-05 evaluation reward: 3.99
episode: 1288 score: 4.0 memory length: 270793 epsilon: 0.6618278800
073414 steps: 298 lr: 4e-05 evaluation reward: 3.97
episode: 1289 score: 5.0 memory length: 271091 epsilon: 0.6612378400
073542 steps: 298 lr: 4e-05 evaluation reward: 4.0
episode: 1290 score: 5.0 memory length: 271377 epsilon: 0.6606715600
073665 steps: 286 lr: 4e-05 evaluation reward: 3.96
episode: 1291 score: 7.0 memory length: 271786 epsilon: 0.6598617400
073841 steps: 409 lr: 4e-05 evaluation reward: 3.99
episode: 1292 score: 3.0 memory length: 272016 epsilon: 0.6594063400
07394 steps: 230 lr: 4e-05 evaluation reward: 3.97
episode: 1293 score: 4.0 memory length: 272275 epsilon: 0.6588935200
074051 steps: 259 lr: 4e-05 evaluation reward: 3.94
```

```
episode: 1294    score: 4.0    memory length: 272535    epsilon: 0.6583787200
074163    steps: 260    lr: 4e-05    evaluation reward: 3.95
episode: 1295    score: 3.0    memory length: 272801    epsilon: 0.6578520400
074277    steps: 266    lr: 4e-05    evaluation reward: 3.95
episode: 1296    score: 5.0    memory length: 273127    epsilon: 0.6572065600
074417    steps: 326    lr: 4e-05    evaluation reward: 3.98
episode: 1297    score: 5.0    memory length: 273433    epsilon: 0.6566006800
074549    steps: 306    lr: 4e-05    evaluation reward: 3.99
episode: 1298    score: 10.0    memory length: 273916    epsilon: 0.655644340
0074756    steps: 483    lr: 4e-05    evaluation reward: 4.05
episode: 1299    score: 3.0    memory length: 274144    epsilon: 0.6551929000
074854    steps: 228    lr: 4e-05    evaluation reward: 4.02
episode: 1300    score: 6.0    memory length: 274508    epsilon: 0.6544721800
075011    steps: 364    lr: 4e-05    evaluation reward: 4.01
episode: 1301    score: 5.0    memory length: 274812    epsilon: 0.6538702600
075141    steps: 304    lr: 4e-05    evaluation reward: 4.06
episode: 1302    score: 3.0    memory length: 275042    epsilon: 0.6534148600
07524    steps: 230    lr: 4e-05    evaluation reward: 4.01
episode: 1303    score: 6.0    memory length: 275401    epsilon: 0.6527040400
075395    steps: 359    lr: 4e-05    evaluation reward: 4.04
episode: 1304    score: 1.0    memory length: 275570    epsilon: 0.6523694200
075467    steps: 169    lr: 4e-05    evaluation reward: 4.02
episode: 1305    score: 6.0    memory length: 275928    epsilon: 0.6516605800
075621    steps: 358    lr: 4e-05    evaluation reward: 4.05
episode: 1306    score: 8.0    memory length: 276369    epsilon: 0.6507874000
075811    steps: 441    lr: 4e-05    evaluation reward: 4.11
episode: 1307    score: 1.0    memory length: 276519    epsilon: 0.6504904000
075875    steps: 150    lr: 4e-05    evaluation reward: 4.09
episode: 1308    score: 6.0    memory length: 276882    epsilon: 0.6497716600
076031    steps: 363    lr: 4e-05    evaluation reward: 4.1
episode: 1309    score: 5.0    memory length: 277190    epsilon: 0.6491618200
076164    steps: 308    lr: 4e-05    evaluation reward: 4.13
episode: 1310    score: 5.0    memory length: 277506    epsilon: 0.6485361400
076299    steps: 316    lr: 4e-05    evaluation reward: 4.16
episode: 1311    score: 3.0    memory length: 277752    epsilon: 0.6480490600
076405    steps: 246    lr: 4e-05    evaluation reward: 4.17
episode: 1312    score: 4.0    memory length: 277993    epsilon: 0.6475718800
076509    steps: 241    lr: 4e-05    evaluation reward: 4.18
episode: 1313    score: 3.0    memory length: 278239    epsilon: 0.6470848000
076614    steps: 246    lr: 4e-05    evaluation reward: 4.21
episode: 1314    score: 2.0    memory length: 278456    epsilon: 0.6466551400
076708    steps: 217    lr: 4e-05    evaluation reward: 4.21
episode: 1315    score: 3.0    memory length: 278702    epsilon: 0.6461680600
076813    steps: 246    lr: 4e-05    evaluation reward: 4.21
episode: 1316    score: 4.0    memory length: 278977    epsilon: 0.6456235600
076932    steps: 275    lr: 4e-05    evaluation reward: 4.22
episode: 1317    score: 3.0    memory length: 279190    epsilon: 0.6452018200
077023    steps: 213    lr: 4e-05    evaluation reward: 4.21
episode: 1318    score: 2.0    memory length: 279388    epsilon: 0.6448097800
077108    steps: 198    lr: 4e-05    evaluation reward: 4.22
episode: 1319    score: 6.0    memory length: 279715    epsilon: 0.6441623200
077249    steps: 327    lr: 4e-05    evaluation reward: 4.25
episode: 1320    score: 2.0    memory length: 279895    epsilon: 0.6438059200
```

```
077326 steps: 180 lr: 4e-05 evaluation reward: 4.17
episode: 1321 score: 4.0 memory length: 280194 epsilon: 0.6432139000
077455 steps: 299 lr: 4e-05 evaluation reward: 4.16
episode: 1322 score: 4.0 memory length: 280464 epsilon: 0.6426793000
077571 steps: 270 lr: 4e-05 evaluation reward: 4.16
episode: 1323 score: 2.0 memory length: 280680 epsilon: 0.6422516200
077664 steps: 216 lr: 4e-05 evaluation reward: 4.16
episode: 1324 score: 5.0 memory length: 281004 epsilon: 0.6416101000
077803 steps: 324 lr: 4e-05 evaluation reward: 4.17
episode: 1325 score: 4.0 memory length: 281299 epsilon: 0.6410260000
07793 steps: 295 lr: 4e-05 evaluation reward: 4.18
episode: 1326 score: 3.0 memory length: 281528 epsilon: 0.6405725800
078028 steps: 229 lr: 4e-05 evaluation reward: 4.19
episode: 1327 score: 9.0 memory length: 281983 epsilon: 0.6396716800
078224 steps: 455 lr: 4e-05 evaluation reward: 4.26
episode: 1328 score: 3.0 memory length: 282251 epsilon: 0.6391410400
078339 steps: 268 lr: 4e-05 evaluation reward: 4.19
episode: 1329 score: 5.0 memory length: 282540 epsilon: 0.6385688200
078463 steps: 289 lr: 4e-05 evaluation reward: 4.17
episode: 1330 score: 5.0 memory length: 282852 epsilon: 0.6379510600
078597 steps: 312 lr: 4e-05 evaluation reward: 4.17
episode: 1331 score: 6.0 memory length: 283205 epsilon: 0.6372521200
078749 steps: 353 lr: 4e-05 evaluation reward: 4.22
episode: 1332 score: 8.0 memory length: 283673 epsilon: 0.6363254800
07895 steps: 468 lr: 4e-05 evaluation reward: 4.26
episode: 1333 score: 3.0 memory length: 283885 epsilon: 0.6359057200
079041 steps: 212 lr: 4e-05 evaluation reward: 4.26
episode: 1334 score: 3.0 memory length: 284132 epsilon: 0.6354166600
079147 steps: 247 lr: 4e-05 evaluation reward: 4.24
episode: 1335 score: 3.0 memory length: 284378 epsilon: 0.6349295800
079253 steps: 246 lr: 4e-05 evaluation reward: 4.23
episode: 1336 score: 1.0 memory length: 284547 epsilon: 0.6345949600
079326 steps: 169 lr: 4e-05 evaluation reward: 4.18
episode: 1337 score: 9.0 memory length: 285012 epsilon: 0.6336742600
079526 steps: 465 lr: 4e-05 evaluation reward: 4.26
episode: 1338 score: 3.0 memory length: 285259 epsilon: 0.6331852000
079632 steps: 247 lr: 4e-05 evaluation reward: 4.27
episode: 1339 score: 1.0 memory length: 285409 epsilon: 0.6328882000
079696 steps: 150 lr: 4e-05 evaluation reward: 4.27
episode: 1340 score: 3.0 memory length: 285655 epsilon: 0.6324011200
079802 steps: 246 lr: 4e-05 evaluation reward: 4.3
episode: 1341 score: 8.0 memory length: 286053 epsilon: 0.6316130800
079973 steps: 398 lr: 4e-05 evaluation reward: 4.31
episode: 1342 score: 4.0 memory length: 286324 epsilon: 0.6310765000
08009 steps: 271 lr: 4e-05 evaluation reward: 4.3
episode: 1343 score: 5.0 memory length: 286647 epsilon: 0.6304369600
080229 steps: 323 lr: 4e-05 evaluation reward: 4.27
episode: 1344 score: 6.0 memory length: 287021 epsilon: 0.6296964400
080389 steps: 374 lr: 4e-05 evaluation reward: 4.3
episode: 1345 score: 5.0 memory length: 287327 epsilon: 0.6290905600
080521 steps: 306 lr: 4e-05 evaluation reward: 4.31
episode: 1346 score: 1.0 memory length: 287478 epsilon: 0.6287915800
080586 steps: 151 lr: 4e-05 evaluation reward: 4.28
```



```
episode: 1347  score: 3.0  memory length: 287703  epsilon: 0.6283460800
080682  steps: 225  lr: 4e-05  evaluation reward: 4.28
episode: 1348  score: 2.0  memory length: 287884  epsilon: 0.6279877000
08076  steps: 181  lr: 4e-05  evaluation reward: 4.23
episode: 1349  score: 2.0  memory length: 288100  epsilon: 0.6275600200
080853  steps: 216  lr: 4e-05  evaluation reward: 4.21
episode: 1350  score: 5.0  memory length: 288424  epsilon: 0.6269185000
080992  steps: 324  lr: 4e-05  evaluation reward: 4.24
episode: 1351  score: 5.0  memory length: 288732  epsilon: 0.6263086600
081125  steps: 308  lr: 4e-05  evaluation reward: 4.23
episode: 1352  score: 3.0  memory length: 288983  epsilon: 0.6258116800
081233  steps: 251  lr: 4e-05  evaluation reward: 4.22
episode: 1353  score: 1.0  memory length: 289153  epsilon: 0.6254750800
081306  steps: 170  lr: 4e-05  evaluation reward: 4.16
episode: 1354  score: 3.0  memory length: 289398  epsilon: 0.6249899800
081411  steps: 245  lr: 4e-05  evaluation reward: 4.16
episode: 1355  score: 8.0  memory length: 289867  epsilon: 0.6240613600
081613  steps: 469  lr: 4e-05  evaluation reward: 4.17
episode: 1356  score: 9.0  memory length: 290375  epsilon: 0.6230555200
081831  steps: 508  lr: 4e-05  evaluation reward: 4.19
episode: 1357  score: 2.0  memory length: 290572  epsilon: 0.6226654600
081916  steps: 197  lr: 4e-05  evaluation reward: 4.18
episode: 1358  score: 3.0  memory length: 290801  epsilon: 0.6222120400
082014  steps: 229  lr: 4e-05  evaluation reward: 4.18
episode: 1359  score: 7.0  memory length: 291207  epsilon: 0.6214081600
082189  steps: 406  lr: 4e-05  evaluation reward: 4.16
episode: 1360  score: 1.0  memory length: 291378  epsilon: 0.6210695800
082262  steps: 171  lr: 4e-05  evaluation reward: 4.12
episode: 1361  score: 1.0  memory length: 291529  epsilon: 0.6207706000
082327  steps: 151  lr: 4e-05  evaluation reward: 4.1
episode: 1362  score: 6.0  memory length: 291885  epsilon: 0.6200657200
08248  steps: 356  lr: 4e-05  evaluation reward: 4.14
episode: 1363  score: 5.0  memory length: 292227  epsilon: 0.6193885600
082627  steps: 342  lr: 4e-05  evaluation reward: 4.15
episode: 1364  score: 3.0  memory length: 292456  epsilon: 0.6189351400
082725  steps: 229  lr: 4e-05  evaluation reward: 4.14
episode: 1365  score: 5.0  memory length: 292763  epsilon: 0.6183272800
082857  steps: 307  lr: 4e-05  evaluation reward: 4.16
episode: 1366  score: 1.0  memory length: 292935  epsilon: 0.6179867200
082931  steps: 172  lr: 4e-05  evaluation reward: 4.12
episode: 1367  score: 5.0  memory length: 293252  epsilon: 0.6173590600
083068  steps: 317  lr: 4e-05  evaluation reward: 4.13
episode: 1368  score: 3.0  memory length: 293519  epsilon: 0.6168304000
083182  steps: 267  lr: 4e-05  evaluation reward: 4.14
episode: 1369  score: 7.0  memory length: 293911  epsilon: 0.6160542400
083351  steps: 392  lr: 4e-05  evaluation reward: 4.18
episode: 1370  score: 3.0  memory length: 294138  epsilon: 0.6156047800
083448  steps: 227  lr: 4e-05  evaluation reward: 4.16
episode: 1371  score: 3.0  memory length: 294368  epsilon: 0.6151493800
083547  steps: 230  lr: 4e-05  evaluation reward: 4.13
episode: 1372  score: 4.0  memory length: 294663  epsilon: 0.6145652800
083674  steps: 295  lr: 4e-05  evaluation reward: 4.05
episode: 1373  score: 3.0  memory length: 294908  epsilon: 0.6140801800
```

```
083779    steps: 245    lr: 4e-05    evaluation reward: 4.06
episode: 1374    score: 6.0    memory length: 295266    epsilon: 0.6133713400
083933    steps: 358    lr: 4e-05    evaluation reward: 4.11
episode: 1375    score: 8.0    memory length: 295728    epsilon: 0.6124565800
084132    steps: 462    lr: 4e-05    evaluation reward: 4.18
episode: 1376    score: 4.0    memory length: 296003    epsilon: 0.6119120800
08425     steps: 275    lr: 4e-05    evaluation reward: 4.19
episode: 1377    score: 5.0    memory length: 296309    epsilon: 0.6113062000
084382    steps: 306    lr: 4e-05    evaluation reward: 4.2
episode: 1378    score: 5.0    memory length: 296654    epsilon: 0.6106231000
08453     steps: 345    lr: 4e-05    evaluation reward: 4.19
episode: 1379    score: 3.0    memory length: 296865    epsilon: 0.6102053200
084621    steps: 211    lr: 4e-05    evaluation reward: 4.19
episode: 1380    score: 6.0    memory length: 297222    epsilon: 0.6094984600
084774    steps: 357    lr: 4e-05    evaluation reward: 4.22
episode: 1381    score: 3.0    memory length: 297470    epsilon: 0.6090074200
084881    steps: 248    lr: 4e-05    evaluation reward: 4.24
episode: 1382    score: 2.0    memory length: 297670    epsilon: 0.6086114200
084967    steps: 200    lr: 4e-05    evaluation reward: 4.22
episode: 1383    score: 5.0    memory length: 297996    epsilon: 0.6079659400
085107    steps: 326    lr: 4e-05    evaluation reward: 4.21
episode: 1384    score: 7.0    memory length: 298244    epsilon: 0.6074749000
085213    steps: 248    lr: 4e-05    evaluation reward: 4.24
episode: 1385    score: 3.0    memory length: 298470    epsilon: 0.6070274200
08531     steps: 226    lr: 4e-05    evaluation reward: 4.23
episode: 1386    score: 4.0    memory length: 298747    epsilon: 0.6064789600
08543     steps: 277    lr: 4e-05    evaluation reward: 4.22
episode: 1387    score: 3.0    memory length: 298997    epsilon: 0.6059839600
085537    steps: 250    lr: 4e-05    evaluation reward: 4.22
episode: 1388    score: 6.0    memory length: 299349    epsilon: 0.6052870000
085688    steps: 352    lr: 4e-05    evaluation reward: 4.24
episode: 1389    score: 1.0    memory length: 299517    epsilon: 0.6049543600
08576     steps: 168    lr: 4e-05    evaluation reward: 4.2
episode: 1390    score: 2.0    memory length: 299716    epsilon: 0.6045603400
085846    steps: 199    lr: 4e-05    evaluation reward: 4.17
episode: 1391    score: 5.0    memory length: 300022    epsilon: 0.6039544600
085978    steps: 306    lr: 1.6000000000000003e-05    evaluation reward:
4.15
episode: 1392    score: 5.0    memory length: 300347    epsilon: 0.6033109600
086117    steps: 325    lr: 1.6000000000000003e-05    evaluation reward:
4.17
episode: 1393    score: 7.0    memory length: 300733    epsilon: 0.6025466800
086283    steps: 386    lr: 1.6000000000000003e-05    evaluation reward:
4.2
episode: 1394    score: 5.0    memory length: 301015    epsilon: 0.6019883200
086404    steps: 282    lr: 1.6000000000000003e-05    evaluation reward:
4.21
episode: 1395    score: 6.0    memory length: 301385    epsilon: 0.6012557200
086563    steps: 370    lr: 1.6000000000000003e-05    evaluation reward:
4.24
episode: 1396    score: 7.0    memory length: 301775    epsilon: 0.6004835200
086731    steps: 390    lr: 1.6000000000000003e-05    evaluation reward:
4.26
```

episode: 1397	score: 4.0	memory length: 302016	epsilon: 0.6000063400
086835	steps: 241	lr: 1.6000000000000003e-05	evaluation reward:
4.25			
episode: 1398	score: 5.0	memory length: 302314	epsilon: 0.5994163000
086963	steps: 298	lr: 1.6000000000000003e-05	evaluation reward:
4.2			
episode: 1399	score: 6.0	memory length: 302665	epsilon: 0.5987213200
087114	steps: 351	lr: 1.6000000000000003e-05	evaluation reward:
4.23			
episode: 1400	score: 4.0	memory length: 302942	epsilon: 0.5981728600
087233	steps: 277	lr: 1.6000000000000003e-05	evaluation reward:
4.21			
episode: 1401	score: 1.0	memory length: 303092	epsilon: 0.5978758600
087297	steps: 150	lr: 1.6000000000000003e-05	evaluation reward:
4.17			
episode: 1402	score: 1.0	memory length: 303243	epsilon: 0.5975768800
087362	steps: 151	lr: 1.6000000000000003e-05	evaluation reward:
4.15			
episode: 1403	score: 4.0	memory length: 303501	epsilon: 0.5970660400
087473	steps: 258	lr: 1.6000000000000003e-05	evaluation reward:
4.13			
episode: 1404	score: 9.0	memory length: 303993	epsilon: 0.5960918800
087684	steps: 492	lr: 1.6000000000000003e-05	evaluation reward:
4.21			
episode: 1405	score: 9.0	memory length: 304486	epsilon: 0.5951157400
087896	steps: 493	lr: 1.6000000000000003e-05	evaluation reward:
4.24			
episode: 1406	score: 4.0	memory length: 304729	epsilon: 0.5946346000
088001	steps: 243	lr: 1.6000000000000003e-05	evaluation reward:
4.2			
episode: 1407	score: 3.0	memory length: 304972	epsilon: 0.5941534600
088105	steps: 243	lr: 1.6000000000000003e-05	evaluation reward:
4.22			
episode: 1408	score: 3.0	memory length: 305220	epsilon: 0.5936624200
088212	steps: 248	lr: 1.6000000000000003e-05	evaluation reward:
4.19			
episode: 1409	score: 4.0	memory length: 305516	epsilon: 0.5930763400
088339	steps: 296	lr: 1.6000000000000003e-05	evaluation reward:
4.18			
episode: 1410	score: 3.0	memory length: 305761	epsilon: 0.5925912400
088444	steps: 245	lr: 1.6000000000000003e-05	evaluation reward:
4.16			
episode: 1411	score: 4.0	memory length: 306018	epsilon: 0.5920823800
088555	steps: 257	lr: 1.6000000000000003e-05	evaluation reward:
4.17			
episode: 1412	score: 5.0	memory length: 306343	epsilon: 0.5914388800
088695	steps: 325	lr: 1.6000000000000003e-05	evaluation reward:
4.18			
episode: 1413	score: 5.0	memory length: 306650	epsilon: 0.5908310200
088827	steps: 307	lr: 1.6000000000000003e-05	evaluation reward:
4.2			
episode: 1414	score: 2.0	memory length: 306848	epsilon: 0.5904389800
088912	steps: 198	lr: 1.6000000000000003e-05	evaluation reward:

```
4.2
episode: 1415  score: 2.0  memory length: 307030  epsilon: 0.5900786200
08899  steps: 182  lr: 1.6000000000000003e-05  evaluation reward: 4
.19
episode: 1416  score: 5.0  memory length: 307368  epsilon: 0.5894093800
089135  steps: 338  lr: 1.6000000000000003e-05  evaluation reward:
4.2
episode: 1417  score: 6.0  memory length: 307683  epsilon: 0.5887856800
089271  steps: 315  lr: 1.6000000000000003e-05  evaluation reward:
4.23
episode: 1418  score: 6.0  memory length: 308046  epsilon: 0.5880669400
089427  steps: 363  lr: 1.6000000000000003e-05  evaluation reward:
4.27
episode: 1419  score: 1.0  memory length: 308197  epsilon: 0.5877679600
089492  steps: 151  lr: 1.6000000000000003e-05  evaluation reward:
4.22
episode: 1420  score: 3.0  memory length: 308410  epsilon: 0.5873462200
089583  steps: 213  lr: 1.6000000000000003e-05  evaluation reward:
4.23
episode: 1421  score: 3.0  memory length: 308658  epsilon: 0.5868551800
08969  steps: 248  lr: 1.6000000000000003e-05  evaluation reward: 4
.22
episode: 1422  score: 1.0  memory length: 308830  epsilon: 0.5865146200
089764  steps: 172  lr: 1.6000000000000003e-05  evaluation reward:
4.19
episode: 1423  score: 6.0  memory length: 309227  epsilon: 0.5857285600
089934  steps: 397  lr: 1.6000000000000003e-05  evaluation reward:
4.23
episode: 1424  score: 7.0  memory length: 309643  epsilon: 0.5849048800
090113  steps: 416  lr: 1.6000000000000003e-05  evaluation reward:
4.25
episode: 1425  score: 6.0  memory length: 309997  epsilon: 0.5842039600
090265  steps: 354  lr: 1.6000000000000003e-05  evaluation reward:
4.27
episode: 1426  score: 3.0  memory length: 310209  epsilon: 0.5837842000
090356  steps: 212  lr: 1.6000000000000003e-05  evaluation reward:
4.27
episode: 1427  score: 5.0  memory length: 310517  epsilon: 0.5831743600
090489  steps: 308  lr: 1.6000000000000003e-05  evaluation reward:
4.23
episode: 1428  score: 6.0  memory length: 310852  epsilon: 0.5825110600
090633  steps: 335  lr: 1.6000000000000003e-05  evaluation reward:
4.26
episode: 1429  score: 3.0  memory length: 311098  epsilon: 0.5820239800
090738  steps: 246  lr: 1.6000000000000003e-05  evaluation reward:
4.24
episode: 1430  score: 4.0  memory length: 311376  epsilon: 0.5814735400
090858  steps: 278  lr: 1.6000000000000003e-05  evaluation reward:
4.23
episode: 1431  score: 2.0  memory length: 311562  epsilon: 0.5811052600
090938  steps: 186  lr: 1.6000000000000003e-05  evaluation reward:
4.19
episode: 1432  score: 1.0  memory length: 311730  epsilon: 0.5807726200
```

09101	steps: 168	lr: 1.6000000000000003e-05	evaluation reward: 4
.12			
episode: 1433	score: 3.0	memory length: 311959	epsilon: 0.5803192000
091109	steps: 229	lr: 1.6000000000000003e-05	evaluation reward:
4.12			
episode: 1434	score: 2.0	memory length: 312141	epsilon: 0.5799588400
091187	steps: 182	lr: 1.6000000000000003e-05	evaluation reward:
4.11			
episode: 1435	score: 6.0	memory length: 312532	epsilon: 0.5791846600
091355	steps: 391	lr: 1.6000000000000003e-05	evaluation reward:
4.14			
episode: 1436	score: 7.0	memory length: 312954	epsilon: 0.5783491000
091536	steps: 422	lr: 1.6000000000000003e-05	evaluation reward:
4.2			
episode: 1437	score: 7.0	memory length: 313367	epsilon: 0.5775313600
091714	steps: 413	lr: 1.6000000000000003e-05	evaluation reward:
4.18			
episode: 1438	score: 3.0	memory length: 313614	epsilon: 0.5770423000
09182	steps: 247	lr: 1.6000000000000003e-05	evaluation reward: 4
.18			
episode: 1439	score: 2.0	memory length: 313797	epsilon: 0.5766799600
091899	steps: 183	lr: 1.6000000000000003e-05	evaluation reward:
4.19			
episode: 1440	score: 3.0	memory length: 314042	epsilon: 0.5761948600
092004	steps: 245	lr: 1.6000000000000003e-05	evaluation reward:
4.19			
episode: 1441	score: 5.0	memory length: 314382	epsilon: 0.5755216600
09215	steps: 340	lr: 1.6000000000000003e-05	evaluation reward: 4
.16			
episode: 1442	score: 5.0	memory length: 314686	epsilon: 0.5749197400
092281	steps: 304	lr: 1.6000000000000003e-05	evaluation reward:
4.17			
episode: 1443	score: 7.0	memory length: 315092	epsilon: 0.5741158600
092455	steps: 406	lr: 1.6000000000000003e-05	evaluation reward:
4.19			
episode: 1444	score: 3.0	memory length: 315317	epsilon: 0.5736703600
092552	steps: 225	lr: 1.6000000000000003e-05	evaluation reward:
4.16			
episode: 1445	score: 3.0	memory length: 315529	epsilon: 0.5732506000
092643	steps: 212	lr: 1.6000000000000003e-05	evaluation reward:
4.14			
episode: 1446	score: 6.0	memory length: 315866	epsilon: 0.5725833400
092788	steps: 337	lr: 1.6000000000000003e-05	evaluation reward:
4.19			
episode: 1447	score: 3.0	memory length: 316092	epsilon: 0.5721358600
092885	steps: 226	lr: 1.6000000000000003e-05	evaluation reward:
4.19			
episode: 1448	score: 3.0	memory length: 316305	epsilon: 0.5717141200
092977	steps: 213	lr: 1.6000000000000003e-05	evaluation reward:
4.2			
episode: 1449	score: 2.0	memory length: 316506	epsilon: 0.5713161400
093063	steps: 201	lr: 1.6000000000000003e-05	evaluation reward:
4.2			

episode: 1450	score: 7.0	memory length: 316910	epsilon: 0.5705162200
093237	steps: 404	lr: 1.6000000000000003e-05	evaluation reward:
4.22			
episode: 1451	score: 2.0	memory length: 317090	epsilon: 0.5701598200
093314	steps: 180	lr: 1.6000000000000003e-05	evaluation reward:
4.19			
episode: 1452	score: 5.0	memory length: 317439	epsilon: 0.5694688000
093464	steps: 349	lr: 1.6000000000000003e-05	evaluation reward:
4.21			
episode: 1453	score: 5.0	memory length: 317767	epsilon: 0.5688193600
093605	steps: 328	lr: 1.6000000000000003e-05	evaluation reward:
4.25			
episode: 1454	score: 4.0	memory length: 318029	epsilon: 0.5683006000
093718	steps: 262	lr: 1.6000000000000003e-05	evaluation reward:
4.26			
episode: 1455	score: 2.0	memory length: 318214	epsilon: 0.5679343000
093797	steps: 185	lr: 1.6000000000000003e-05	evaluation reward:
4.2			
episode: 1456	score: 5.0	memory length: 318539	epsilon: 0.5672908000
093937	steps: 325	lr: 1.6000000000000003e-05	evaluation reward:
4.16			
episode: 1457	score: 7.0	memory length: 318966	epsilon: 0.5664453400
09412	steps: 427	lr: 1.6000000000000003e-05	evaluation reward: 4
.21			
episode: 1458	score: 4.0	memory length: 319208	epsilon: 0.5659661800
094224	steps: 242	lr: 1.6000000000000003e-05	evaluation reward:
4.22			
episode: 1459	score: 4.0	memory length: 319505	epsilon: 0.5653781200
094352	steps: 297	lr: 1.6000000000000003e-05	evaluation reward:
4.19			
episode: 1460	score: 3.0	memory length: 319734	epsilon: 0.5649247000
094451	steps: 229	lr: 1.6000000000000003e-05	evaluation reward:
4.21			
episode: 1461	score: 3.0	memory length: 319985	epsilon: 0.5644277200
094558	steps: 251	lr: 1.6000000000000003e-05	evaluation reward:
4.23			
episode: 1462	score: 3.0	memory length: 320211	epsilon: 0.5639802400
094656	steps: 226	lr: 1.6000000000000003e-05	evaluation reward:
4.2			
episode: 1463	score: 2.0	memory length: 320393	epsilon: 0.5636198800
094734	steps: 182	lr: 1.6000000000000003e-05	evaluation reward:
4.17			
episode: 1464	score: 7.0	memory length: 320789	epsilon: 0.5628358000
094904	steps: 396	lr: 1.6000000000000003e-05	evaluation reward:
4.21			
episode: 1465	score: 7.0	memory length: 321142	epsilon: 0.5621368600
095056	steps: 353	lr: 1.6000000000000003e-05	evaluation reward:
4.23			
episode: 1466	score: 3.0	memory length: 321371	epsilon: 0.5616834400
095154	steps: 229	lr: 1.6000000000000003e-05	evaluation reward:
4.25			
episode: 1467	score: 4.0	memory length: 321636	epsilon: 0.5611587400
095268	steps: 265	lr: 1.6000000000000003e-05	evaluation reward:

4.24
episode: 1468 score: 9.0 memory length: 322056 epsilon: 0.5603271400
095449 steps: 420 lr: 1.6000000000000003e-05 evaluation reward:
4.3
episode: 1469 score: 4.0 memory length: 322332 epsilon: 0.5597806600
095567 steps: 276 lr: 1.6000000000000003e-05 evaluation reward:
4.27
episode: 1470 score: 4.0 memory length: 322625 epsilon: 0.5592005200
095693 steps: 293 lr: 1.6000000000000003e-05 evaluation reward:
4.28
episode: 1471 score: 3.0 memory length: 322852 epsilon: 0.5587510600
095791 steps: 227 lr: 1.6000000000000003e-05 evaluation reward:
4.28
episode: 1472 score: 4.0 memory length: 323128 epsilon: 0.5582045800
09591 steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 4
.28
episode: 1473 score: 7.0 memory length: 323374 epsilon: 0.5577175000
096015 steps: 246 lr: 1.6000000000000003e-05 evaluation reward:
4.32
episode: 1474 score: 3.0 memory length: 323621 epsilon: 0.5572284400
096121 steps: 247 lr: 1.6000000000000003e-05 evaluation reward:
4.29
episode: 1475 score: 7.0 memory length: 324024 epsilon: 0.5564305000
096295 steps: 403 lr: 1.6000000000000003e-05 evaluation reward:
4.28
episode: 1476 score: 3.0 memory length: 324271 epsilon: 0.5559414400
096401 steps: 247 lr: 1.6000000000000003e-05 evaluation reward:
4.27
episode: 1477 score: 4.0 memory length: 324567 epsilon: 0.5553553600
096528 steps: 296 lr: 1.6000000000000003e-05 evaluation reward:
4.26
episode: 1478 score: 6.0 memory length: 324940 epsilon: 0.5546168200
096688 steps: 373 lr: 1.6000000000000003e-05 evaluation reward:
4.27
episode: 1479 score: 4.0 memory length: 325199 epsilon: 0.5541040000
0968 steps: 259 lr: 1.6000000000000003e-05 evaluation reward: 4.
28
episode: 1480 score: 3.0 memory length: 325444 epsilon: 0.5536189000
096905 steps: 245 lr: 1.6000000000000003e-05 evaluation reward:
4.25
episode: 1481 score: 3.0 memory length: 325670 epsilon: 0.5531714200
097002 steps: 226 lr: 1.6000000000000003e-05 evaluation reward:
4.25
episode: 1482 score: 2.0 memory length: 325868 epsilon: 0.5527793800
097087 steps: 198 lr: 1.6000000000000003e-05 evaluation reward:
4.25
episode: 1483 score: 13.0 memory length: 326356 epsilon: 0.551813140
0097297 steps: 488 lr: 1.6000000000000003e-05 evaluation reward:
4.33
episode: 1484 score: 6.0 memory length: 326719 epsilon: 0.5510944000
097453 steps: 363 lr: 1.6000000000000003e-05 evaluation reward:
4.32
episode: 1485 score: 5.0 memory length: 327044 epsilon: 0.5504509000

097593	steps: 325	lr: 1.6000000000000003e-05	evaluation reward:
4.34			
episode: 1486	score: 2.0	memory length: 327224	epsilon: 0.5500945000
09767	steps: 180	lr: 1.6000000000000003e-05	evaluation reward: 4
.32			
episode: 1487	score: 9.0	memory length: 327565	epsilon: 0.5494193200
097817	steps: 341	lr: 1.6000000000000003e-05	evaluation reward:
4.38			
episode: 1488	score: 5.0	memory length: 327913	epsilon: 0.5487302800
097966	steps: 348	lr: 1.6000000000000003e-05	evaluation reward:
4.37			
episode: 1489	score: 8.0	memory length: 328356	epsilon: 0.5478531400
098157	steps: 443	lr: 1.6000000000000003e-05	evaluation reward:
4.44			
episode: 1490	score: 5.0	memory length: 328702	epsilon: 0.5471680600
098305	steps: 346	lr: 1.6000000000000003e-05	evaluation reward:
4.47			
episode: 1491	score: 4.0	memory length: 328977	epsilon: 0.5466235600
098424	steps: 275	lr: 1.6000000000000003e-05	evaluation reward:
4.46			
episode: 1492	score: 2.0	memory length: 329158	epsilon: 0.5462651800
098501	steps: 181	lr: 1.6000000000000003e-05	evaluation reward:
4.43			
episode: 1493	score: 3.0	memory length: 329386	epsilon: 0.5458137400
098599	steps: 228	lr: 1.6000000000000003e-05	evaluation reward:
4.39			
episode: 1494	score: 3.0	memory length: 329612	epsilon: 0.5453662600
098697	steps: 226	lr: 1.6000000000000003e-05	evaluation reward:
4.37			
episode: 1495	score: 8.0	memory length: 330064	epsilon: 0.5444713000
098891	steps: 452	lr: 1.6000000000000003e-05	evaluation reward:
4.39			
episode: 1496	score: 3.0	memory length: 330312	epsilon: 0.5439802600
098997	steps: 248	lr: 1.6000000000000003e-05	evaluation reward:
4.35			
episode: 1497	score: 5.0	memory length: 330617	epsilon: 0.5433763600
099129	steps: 305	lr: 1.6000000000000003e-05	evaluation reward:
4.36			
episode: 1498	score: 10.0	memory length: 331124	epsilon: 0.542372500
0099346	steps: 507	lr: 1.6000000000000003e-05	evaluation reward:
4.41			
episode: 1499	score: 4.0	memory length: 331399	epsilon: 0.5418280000
099465	steps: 275	lr: 1.6000000000000003e-05	evaluation reward:
4.39			
episode: 1500	score: 5.0	memory length: 331703	epsilon: 0.5412260800
099595	steps: 304	lr: 1.6000000000000003e-05	evaluation reward:
4.4			
episode: 1501	score: 1.0	memory length: 331873	epsilon: 0.5408894800
099668	steps: 170	lr: 1.6000000000000003e-05	evaluation reward:
4.4			
episode: 1502	score: 6.0	memory length: 332246	epsilon: 0.5401509400
099829	steps: 373	lr: 1.6000000000000003e-05	evaluation reward:
4.45			

episode: 1503	score: 4.0	memory length: 332507	epsilon: 0.5396341600
099941	steps: 261	lr: 1.6000000000000003e-05	evaluation reward:
4.45			
episode: 1504	score: 6.0	memory length: 332837	epsilon: 0.5389807600
100083	steps: 330	lr: 1.6000000000000003e-05	evaluation reward:
4.42			
episode: 1505	score: 7.0	memory length: 333263	epsilon: 0.5381372800
100266	steps: 426	lr: 1.6000000000000003e-05	evaluation reward:
4.4			
episode: 1506	score: 4.0	memory length: 333559	epsilon: 0.5375512000
100393	steps: 296	lr: 1.6000000000000003e-05	evaluation reward:
4.4			
episode: 1507	score: 7.0	memory length: 333962	epsilon: 0.5367532600
100566	steps: 403	lr: 1.6000000000000003e-05	evaluation reward:
4.44			
episode: 1508	score: 11.0	memory length: 334541	epsilon: 0.535606840
0100815	steps: 579	lr: 1.6000000000000003e-05	evaluation reward:
4.52			
episode: 1509	score: 7.0	memory length: 334939	epsilon: 0.5348188000
100986	steps: 398	lr: 1.6000000000000003e-05	evaluation reward:
4.55			
episode: 1510	score: 6.0	memory length: 335274	epsilon: 0.5341555000
10113	steps: 335	lr: 1.6000000000000003e-05	evaluation reward: 4
.58			
episode: 1511	score: 6.0	memory length: 335623	epsilon: 0.5334644800
10128	steps: 349	lr: 1.6000000000000003e-05	evaluation reward: 4
.6			
episode: 1512	score: 6.0	memory length: 335968	epsilon: 0.5327813800
101429	steps: 345	lr: 1.6000000000000003e-05	evaluation reward:
4.61			
episode: 1513	score: 4.0	memory length: 336225	epsilon: 0.5322725200
101539	steps: 257	lr: 1.6000000000000003e-05	evaluation reward:
4.6			
episode: 1514	score: 3.0	memory length: 336438	epsilon: 0.5318507800
101631	steps: 213	lr: 1.6000000000000003e-05	evaluation reward:
4.61			
episode: 1515	score: 11.0	memory length: 336997	epsilon: 0.530743960
0101871	steps: 559	lr: 1.6000000000000003e-05	evaluation reward:
4.7			
episode: 1516	score: 3.0	memory length: 337222	epsilon: 0.5302984600
101968	steps: 225	lr: 1.6000000000000003e-05	evaluation reward:
4.68			
episode: 1517	score: 2.0	memory length: 337420	epsilon: 0.5299064200
102053	steps: 198	lr: 1.6000000000000003e-05	evaluation reward:
4.64			
episode: 1518	score: 6.0	memory length: 337756	epsilon: 0.5292411400
102197	steps: 336	lr: 1.6000000000000003e-05	evaluation reward:
4.64			
episode: 1519	score: 6.0	memory length: 338065	epsilon: 0.5286293200
10233	steps: 309	lr: 1.6000000000000003e-05	evaluation reward: 4
.69			
episode: 1520	score: 4.0	memory length: 338327	epsilon: 0.5281105600
102443	steps: 262	lr: 1.6000000000000003e-05	evaluation reward:

4.7
episode: 1521 score: 4.0 memory length: 338608 epsilon: 0.5275541800
102563 steps: 281 lr: 1.6000000000000003e-05 evaluation reward:
4.71
episode: 1522 score: 5.0 memory length: 338932 epsilon: 0.5269126600
102703 steps: 324 lr: 1.6000000000000003e-05 evaluation reward:
4.75
episode: 1523 score: 3.0 memory length: 339163 epsilon: 0.5264552800
102802 steps: 231 lr: 1.6000000000000003e-05 evaluation reward:
4.72
episode: 1524 score: 5.0 memory length: 339506 epsilon: 0.5257761400
102949 steps: 343 lr: 1.6000000000000003e-05 evaluation reward:
4.7
episode: 1525 score: 6.0 memory length: 339878 epsilon: 0.5250395800
103109 steps: 372 lr: 1.6000000000000003e-05 evaluation reward:
4.7
episode: 1526 score: 8.0 memory length: 340344 epsilon: 0.5241169000
10331 steps: 466 lr: 1.6000000000000003e-05 evaluation reward: 4
.75
episode: 1527 score: 6.0 memory length: 340697 epsilon: 0.5234179600
103461 steps: 353 lr: 1.6000000000000003e-05 evaluation reward:
4.76
episode: 1528 score: 11.0 memory length: 341163 epsilon: 0.522495280
0103662 steps: 466 lr: 1.6000000000000003e-05 evaluation reward:
4.81
episode: 1529 score: 3.0 memory length: 341389 epsilon: 0.5220478000
103759 steps: 226 lr: 1.6000000000000003e-05 evaluation reward:
4.81
episode: 1530 score: 7.0 memory length: 341772 epsilon: 0.5212894600
103923 steps: 383 lr: 1.6000000000000003e-05 evaluation reward:
4.84
episode: 1531 score: 7.0 memory length: 342208 epsilon: 0.5204261800
104111 steps: 436 lr: 1.6000000000000003e-05 evaluation reward:
4.89
episode: 1532 score: 10.0 memory length: 342726 epsilon: 0.519400540
0104333 steps: 518 lr: 1.6000000000000003e-05 evaluation reward:
4.98
episode: 1533 score: 9.0 memory length: 343243 epsilon: 0.5183768800
104556 steps: 517 lr: 1.6000000000000003e-05 evaluation reward:
5.04
episode: 1534 score: 6.0 memory length: 343634 epsilon: 0.5176027000
104724 steps: 391 lr: 1.6000000000000003e-05 evaluation reward:
5.08
episode: 1535 score: 5.0 memory length: 343958 epsilon: 0.5169611800
104863 steps: 324 lr: 1.6000000000000003e-05 evaluation reward:
5.07
episode: 1536 score: 4.0 memory length: 344236 epsilon: 0.5164107400
104982 steps: 278 lr: 1.6000000000000003e-05 evaluation reward:
5.04
episode: 1537 score: 5.0 memory length: 344514 epsilon: 0.5158603000
105102 steps: 278 lr: 1.6000000000000003e-05 evaluation reward:
5.02
episode: 1538 score: 4.0 memory length: 344791 epsilon: 0.5153118400

105221	steps: 277	lr: 1.6000000000000003e-05	evaluation reward:
5.03			
episode: 1539	score: 8.0	memory length: 345232	epsilon: 0.5144386600
105411	steps: 441	lr: 1.6000000000000003e-05	evaluation reward:
5.09			
episode: 1540	score: 4.0	memory length: 345492	epsilon: 0.5139238600
105522	steps: 260	lr: 1.6000000000000003e-05	evaluation reward:
5.1			
episode: 1541	score: 5.0	memory length: 345817	epsilon: 0.5132803600
105662	steps: 325	lr: 1.6000000000000003e-05	evaluation reward:
5.1			
episode: 1542	score: 4.0	memory length: 346074	epsilon: 0.5127715000
105773	steps: 257	lr: 1.6000000000000003e-05	evaluation reward:
5.09			
episode: 1543	score: 6.0	memory length: 346439	epsilon: 0.5120488000
105929	steps: 365	lr: 1.6000000000000003e-05	evaluation reward:
5.08			
episode: 1544	score: 4.0	memory length: 346696	epsilon: 0.5115399400
10604	steps: 257	lr: 1.6000000000000003e-05	evaluation reward: 5
.09			
episode: 1545	score: 6.0	memory length: 347067	epsilon: 0.5108053600
106199	steps: 371	lr: 1.6000000000000003e-05	evaluation reward:
5.12			
episode: 1546	score: 5.0	memory length: 347411	epsilon: 0.5101242400
106347	steps: 344	lr: 1.6000000000000003e-05	evaluation reward:
5.11			
episode: 1547	score: 9.0	memory length: 347818	epsilon: 0.5093183800
106522	steps: 407	lr: 1.6000000000000003e-05	evaluation reward:
5.17			
episode: 1548	score: 4.0	memory length: 348092	epsilon: 0.5087758600
10664	steps: 274	lr: 1.6000000000000003e-05	evaluation reward: 5
.18			
episode: 1549	score: 6.0	memory length: 348440	epsilon: 0.5080868200
10679	steps: 348	lr: 1.6000000000000003e-05	evaluation reward: 5
.22			
episode: 1550	score: 8.0	memory length: 348878	epsilon: 0.5072195800
106978	steps: 438	lr: 1.6000000000000003e-05	evaluation reward:
5.23			
episode: 1551	score: 11.0	memory length: 349306	epsilon: 0.506372140
0107162	steps: 428	lr: 1.6000000000000003e-05	evaluation reward:
5.32			
episode: 1552	score: 3.0	memory length: 349535	epsilon: 0.5059187200
10726	steps: 229	lr: 1.6000000000000003e-05	evaluation reward: 5
.3			
episode: 1553	score: 7.0	memory length: 349937	epsilon: 0.5051227600
107433	steps: 402	lr: 1.6000000000000003e-05	evaluation reward:
5.32			
episode: 1554	score: 9.0	memory length: 350401	epsilon: 0.5042040400
107632	steps: 464	lr: 1.6000000000000003e-05	evaluation reward:
5.37			
episode: 1555	score: 3.0	memory length: 350647	epsilon: 0.5037169600
107738	steps: 246	lr: 1.6000000000000003e-05	evaluation reward:
5.38			

episode: 1556	score: 7.0	memory length: 351031	epsilon: 0.5029566400
107903	steps: 384	lr: 1.6000000000000003e-05	evaluation reward:
5.4			
episode: 1557	score: 7.0	memory length: 351413	epsilon: 0.5022002800
108067	steps: 382	lr: 1.6000000000000003e-05	evaluation reward:
5.4			
episode: 1558	score: 3.0	memory length: 351638	epsilon: 0.5017547800
108164	steps: 225	lr: 1.6000000000000003e-05	evaluation reward:
5.39			
episode: 1559	score: 4.0	memory length: 351936	epsilon: 0.5011647400
108292	steps: 298	lr: 1.6000000000000003e-05	evaluation reward:
5.39			
episode: 1560	score: 6.0	memory length: 352300	epsilon: 0.5004440200
108449	steps: 364	lr: 1.6000000000000003e-05	evaluation reward:
5.42			
episode: 1561	score: 3.0	memory length: 352510	epsilon: 0.5000282200
108539	steps: 210	lr: 1.6000000000000003e-05	evaluation reward:
5.42			
episode: 1562	score: 3.0	memory length: 352738	epsilon: 0.4995767800
108518	steps: 228	lr: 1.6000000000000003e-05	evaluation reward:
5.42			
episode: 1563	score: 6.0	memory length: 353094	epsilon: 0.4988719000
1084736	steps: 356	lr: 1.6000000000000003e-05	evaluation reward:
5.46			
episode: 1564	score: 1.0	memory length: 353244	epsilon: 0.4985749000
108455	steps: 150	lr: 1.6000000000000003e-05	evaluation reward:
5.4			
episode: 1565	score: 4.0	memory length: 353487	epsilon: 0.4980937600
1084243	steps: 243	lr: 1.6000000000000003e-05	evaluation reward:
5.37			
episode: 1566	score: 2.0	memory length: 353685	epsilon: 0.4977017200
1083995	steps: 198	lr: 1.6000000000000003e-05	evaluation reward:
5.36			
episode: 1567	score: 5.0	memory length: 353994	epsilon: 0.4970899000
108361	steps: 309	lr: 1.6000000000000003e-05	evaluation reward:
5.37			
episode: 1568	score: 6.0	memory length: 354369	epsilon: 0.4963474000
108314	steps: 375	lr: 1.6000000000000003e-05	evaluation reward:
5.34			
episode: 1569	score: 3.0	memory length: 354616	epsilon: 0.4958583400
108283	steps: 247	lr: 1.6000000000000003e-05	evaluation reward:
5.33			
episode: 1570	score: 6.0	memory length: 354938	epsilon: 0.4952207800
1082426	steps: 322	lr: 1.6000000000000003e-05	evaluation reward:
5.35			
episode: 1571	score: 4.0	memory length: 355234	epsilon: 0.4946347000
1082055	steps: 296	lr: 1.6000000000000003e-05	evaluation reward:
5.36			
episode: 1572	score: 6.0	memory length: 355609	epsilon: 0.4938922000
1081585	steps: 375	lr: 1.6000000000000003e-05	evaluation reward:
5.38			
episode: 1573	score: 3.0	memory length: 355857	epsilon: 0.4934011600
1081274	steps: 248	lr: 1.6000000000000003e-05	evaluation reward:

5.34
episode: 1574 score: 5.0 memory length: 356163 epsilon: 0.4927952800
108089 steps: 306 lr: 1.6000000000000003e-05 evaluation reward:
5.36
episode: 1575 score: 2.0 memory length: 356382 epsilon: 0.4923616600
1080617 steps: 219 lr: 1.6000000000000003e-05 evaluation reward:
5.31
episode: 1576 score: 5.0 memory length: 356685 epsilon: 0.4917617200
108024 steps: 303 lr: 1.6000000000000003e-05 evaluation reward:
5.33
episode: 1577 score: 3.0 memory length: 356914 epsilon: 0.4913083000
107995 steps: 229 lr: 1.6000000000000003e-05 evaluation reward:
5.32
episode: 1578 score: 6.0 memory length: 357285 epsilon: 0.4905737200
1079486 steps: 371 lr: 1.6000000000000003e-05 evaluation reward:
5.32
episode: 1579 score: 8.0 memory length: 357737 epsilon: 0.4896787600
107892 steps: 452 lr: 1.6000000000000003e-05 evaluation reward:
5.36
episode: 1580 score: 2.0 memory length: 357917 epsilon: 0.4893223600
1078694 steps: 180 lr: 1.6000000000000003e-05 evaluation reward:
5.35
episode: 1581 score: 3.0 memory length: 358164 epsilon: 0.4888333000
1078384 steps: 247 lr: 1.6000000000000003e-05 evaluation reward:
5.35
episode: 1582 score: 3.0 memory length: 358430 epsilon: 0.4883066200
107805 steps: 266 lr: 1.6000000000000003e-05 evaluation reward:
5.36
episode: 1583 score: 5.0 memory length: 358756 epsilon: 0.4876611400
1077643 steps: 326 lr: 1.6000000000000003e-05 evaluation reward:
5.28
episode: 1584 score: 8.0 memory length: 359215 epsilon: 0.4867523200
107707 steps: 459 lr: 1.6000000000000003e-05 evaluation reward:
5.3
episode: 1585 score: 4.0 memory length: 359475 epsilon: 0.4862375200
107674 steps: 260 lr: 1.6000000000000003e-05 evaluation reward:
5.29
episode: 1586 score: 1.0 memory length: 359626 epsilon: 0.4859385400
1076553 steps: 151 lr: 1.6000000000000003e-05 evaluation reward:
5.28
episode: 1587 score: 3.0 memory length: 359838 epsilon: 0.4855187800
107629 steps: 212 lr: 1.6000000000000003e-05 evaluation reward:
5.22
episode: 1588 score: 11.0 memory length: 360379 epsilon: 0.484447600
0107561 steps: 541 lr: 1.6000000000000003e-05 evaluation reward:
5.28
episode: 1589 score: 7.0 memory length: 360770 epsilon: 0.4836734200
107512 steps: 391 lr: 1.6000000000000003e-05 evaluation reward:
5.27
episode: 1590 score: 3.0 memory length: 361017 epsilon: 0.4831843600
107481 steps: 247 lr: 1.6000000000000003e-05 evaluation reward:
5.25
episode: 1591 score: 6.0 memory length: 361376 epsilon: 0.4824735400

107436	steps: 359	lr: 1.6000000000000003e-05	evaluation reward:
5.27			
episode: 1592	score: 8.0	memory length: 361826	epsilon: 0.4815825400
1073797	steps: 450	lr: 1.6000000000000003e-05	evaluation reward:
5.33			
episode: 1593	score: 6.0	memory length: 362162	epsilon: 0.4809172600
1073376	steps: 336	lr: 1.6000000000000003e-05	evaluation reward:
5.36			
episode: 1594	score: 6.0	memory length: 362490	epsilon: 0.4802678200
1072965	steps: 328	lr: 1.6000000000000003e-05	evaluation reward:
5.39			
episode: 1595	score: 14.0	memory length: 362877	epsilon: 0.479501560
0107248	steps: 387	lr: 1.6000000000000003e-05	evaluation reward:
5.45			
episode: 1596	score: 2.0	memory length: 363058	epsilon: 0.4791431800
1072254	steps: 181	lr: 1.6000000000000003e-05	evaluation reward:
5.44			
episode: 1597	score: 7.0	memory length: 363471	epsilon: 0.4783254400
1071736	steps: 413	lr: 1.6000000000000003e-05	evaluation reward:
5.46			
episode: 1598	score: 7.0	memory length: 363830	epsilon: 0.4776146200
1071286	steps: 359	lr: 1.6000000000000003e-05	evaluation reward:
5.43			
episode: 1599	score: 3.0	memory length: 364076	epsilon: 0.4771275400
107098	steps: 246	lr: 1.6000000000000003e-05	evaluation reward:
5.42			
episode: 1600	score: 3.0	memory length: 364287	epsilon: 0.4767097600
1070714	steps: 211	lr: 1.6000000000000003e-05	evaluation reward:
5.4			
episode: 1601	score: 8.0	memory length: 364712	epsilon: 0.4758682600
107018	steps: 425	lr: 1.6000000000000003e-05	evaluation reward:
5.47			
episode: 1602	score: 7.0	memory length: 365112	epsilon: 0.4750762600
106968	steps: 400	lr: 1.6000000000000003e-05	evaluation reward:
5.48			
episode: 1603	score: 5.0	memory length: 365439	epsilon: 0.4744288000
106927	steps: 327	lr: 1.6000000000000003e-05	evaluation reward:
5.49			
episode: 1604	score: 5.0	memory length: 365747	epsilon: 0.4738189600
1068885	steps: 308	lr: 1.6000000000000003e-05	evaluation reward:
5.48			
episode: 1605	score: 8.0	memory length: 366217	epsilon: 0.4728883600
1068296	steps: 470	lr: 1.6000000000000003e-05	evaluation reward:
5.49			
episode: 1606	score: 3.0	memory length: 366446	epsilon: 0.4724349400
106801	steps: 229	lr: 1.6000000000000003e-05	evaluation reward:
5.48			
episode: 1607	score: 9.0	memory length: 366941	epsilon: 0.4714548400
106739	steps: 495	lr: 1.6000000000000003e-05	evaluation reward:
5.5			
episode: 1608	score: 6.0	memory length: 367256	epsilon: 0.4708311400
1066995	steps: 315	lr: 1.6000000000000003e-05	evaluation reward:
5.45			

episode: 1609	score: 5.0	memory length: 367582	epsilon: 0.4701856600
1066586	steps: 326	lr: 1.6000000000000003e-05	evaluation reward: 5.43
episode: 1610	score: 3.0	memory length: 367828	epsilon: 0.4696985800
106628	steps: 246	lr: 1.6000000000000003e-05	evaluation reward: 5.4
episode: 1611	score: 3.0	memory length: 368053	epsilon: 0.4692530800
1065996	steps: 225	lr: 1.6000000000000003e-05	evaluation reward: 5.37
episode: 1612	score: 7.0	memory length: 368465	epsilon: 0.4684373200
106548	steps: 412	lr: 1.6000000000000003e-05	evaluation reward: 5.38
episode: 1613	score: 3.0	memory length: 368694	epsilon: 0.4679839000
1065193	steps: 229	lr: 1.6000000000000003e-05	evaluation reward: 5.37
episode: 1614	score: 6.0	memory length: 369023	epsilon: 0.4673324800
106478	steps: 329	lr: 1.6000000000000003e-05	evaluation reward: 5.4
episode: 1615	score: 6.0	memory length: 369417	epsilon: 0.4665523600
106429	steps: 394	lr: 1.6000000000000003e-05	evaluation reward: 5.35
episode: 1616	score: 3.0	memory length: 369627	epsilon: 0.4661365600
1064024	steps: 210	lr: 1.6000000000000003e-05	evaluation reward: 5.35
episode: 1617	score: 5.0	memory length: 369914	epsilon: 0.4655683000
1063665	steps: 287	lr: 1.6000000000000003e-05	evaluation reward: 5.38
episode: 1618	score: 5.0	memory length: 370236	epsilon: 0.4649307400
106326	steps: 322	lr: 1.6000000000000003e-05	evaluation reward: 5.37
episode: 1619	score: 7.0	memory length: 370672	epsilon: 0.4640674600
1062715	steps: 436	lr: 1.6000000000000003e-05	evaluation reward: 5.38
episode: 1620	score: 6.0	memory length: 371007	epsilon: 0.4634041600
1062296	steps: 335	lr: 1.6000000000000003e-05	evaluation reward: 5.4
episode: 1621	score: 9.0	memory length: 371481	epsilon: 0.4624656400
10617	steps: 474	lr: 1.6000000000000003e-05	evaluation reward: 5.45
episode: 1622	score: 7.0	memory length: 371909	epsilon: 0.4616182000
1061166	steps: 428	lr: 1.6000000000000003e-05	evaluation reward: 5.47
episode: 1623	score: 4.0	memory length: 372203	epsilon: 0.4610360800
10608	steps: 294	lr: 1.6000000000000003e-05	evaluation reward: 5.48
episode: 1624	score: 3.0	memory length: 372431	epsilon: 0.4605846400
106051	steps: 228	lr: 1.6000000000000003e-05	evaluation reward: 5.46
episode: 1625	score: 3.0	memory length: 372644	epsilon: 0.4601629000
1060245	steps: 213	lr: 1.6000000000000003e-05	evaluation reward: 5.43
episode: 1626	score: 4.0	memory length: 372924	epsilon: 0.4596085000
1059894	steps: 280	lr: 1.6000000000000003e-05	evaluation reward:

5.39
episode: 1627 score: 3.0 memory length: 373152 epsilon: 0.4591570600
105961 steps: 228 lr: 1.6000000000000003e-05 evaluation reward:
5.36
episode: 1628 score: 7.0 memory length: 373530 epsilon: 0.4584086200
1059135 steps: 378 lr: 1.6000000000000003e-05 evaluation reward:
5.32
episode: 1629 score: 3.0 memory length: 373742 epsilon: 0.4579888600
105887 steps: 212 lr: 1.6000000000000003e-05 evaluation reward:
5.32
episode: 1630 score: 5.0 memory length: 374047 epsilon: 0.4573849600
105849 steps: 305 lr: 1.6000000000000003e-05 evaluation reward:
5.3
episode: 1631 score: 8.0 memory length: 374496 epsilon: 0.4564959400
1057925 steps: 449 lr: 1.6000000000000003e-05 evaluation reward:
5.31
episode: 1632 score: 5.0 memory length: 374804 epsilon: 0.4558861000
105754 steps: 308 lr: 1.6000000000000003e-05 evaluation reward:
5.26
episode: 1633 score: 4.0 memory length: 375042 epsilon: 0.4554148600
105724 steps: 238 lr: 1.6000000000000003e-05 evaluation reward:
5.21
episode: 1634 score: 3.0 memory length: 375270 epsilon: 0.4549634200
1056955 steps: 228 lr: 1.6000000000000003e-05 evaluation reward:
5.18
episode: 1635 score: 5.0 memory length: 375594 epsilon: 0.4543219000
105655 steps: 324 lr: 1.6000000000000003e-05 evaluation reward:
5.18
episode: 1636 score: 9.0 memory length: 376050 epsilon: 0.4534190200
105598 steps: 456 lr: 1.6000000000000003e-05 evaluation reward:
5.23
episode: 1637 score: 4.0 memory length: 376323 epsilon: 0.4528784800
1055636 steps: 273 lr: 1.6000000000000003e-05 evaluation reward:
5.22
episode: 1638 score: 6.0 memory length: 376658 epsilon: 0.4522151800
1055216 steps: 335 lr: 1.6000000000000003e-05 evaluation reward:
5.24
episode: 1639 score: 6.0 memory length: 377025 epsilon: 0.4514885200
1054757 steps: 367 lr: 1.6000000000000003e-05 evaluation reward:
5.22
episode: 1640 score: 3.0 memory length: 377253 epsilon: 0.4510370800
105447 steps: 228 lr: 1.6000000000000003e-05 evaluation reward:
5.21
episode: 1641 score: 5.0 memory length: 377543 epsilon: 0.4504628800
105411 steps: 290 lr: 1.6000000000000003e-05 evaluation reward:
5.21
episode: 1642 score: 4.0 memory length: 377838 epsilon: 0.4498787800
105374 steps: 295 lr: 1.6000000000000003e-05 evaluation reward:
5.21
episode: 1643 score: 8.0 memory length: 378258 epsilon: 0.4490471800
105321 steps: 420 lr: 1.6000000000000003e-05 evaluation reward:
5.23
episode: 1644 score: 7.0 memory length: 378623 epsilon: 0.4483244800

1052755	steps: 365	lr: 1.6000000000000003e-05	evaluation reward:
5.26			
episode: 1645	score: 3.0	memory length: 378870	epsilon: 0.4478354200
1052445	steps: 247	lr: 1.6000000000000003e-05	evaluation reward:
5.23			
episode: 1646	score: 3.0	memory length: 379079	epsilon: 0.4474216000
1052184	steps: 209	lr: 1.6000000000000003e-05	evaluation reward:
5.21			
episode: 1647	score: 8.0	memory length: 379523	epsilon: 0.4465424800
105163	steps: 444	lr: 1.6000000000000003e-05	evaluation reward:
5.2			
episode: 1648	score: 3.0	memory length: 379733	epsilon: 0.4461266800
1051364	steps: 210	lr: 1.6000000000000003e-05	evaluation reward:
5.19			
episode: 1649	score: 6.0	memory length: 380106	epsilon: 0.4453881400
1050897	steps: 373	lr: 1.6000000000000003e-05	evaluation reward:
5.19			
episode: 1650	score: 6.0	memory length: 380452	epsilon: 0.4447030600
1050464	steps: 346	lr: 1.6000000000000003e-05	evaluation reward:
5.17			
episode: 1651	score: 4.0	memory length: 380747	epsilon: 0.4441189600
1050094	steps: 295	lr: 1.6000000000000003e-05	evaluation reward:
5.1			
episode: 1652	score: 3.0	memory length: 380973	epsilon: 0.4436714800
104981	steps: 226	lr: 1.6000000000000003e-05	evaluation reward:
5.1			
episode: 1653	score: 3.0	memory length: 381183	epsilon: 0.4432556800
104955	steps: 210	lr: 1.6000000000000003e-05	evaluation reward:
5.06			
episode: 1654	score: 5.0	memory length: 381500	epsilon: 0.4426280200
104915	steps: 317	lr: 1.6000000000000003e-05	evaluation reward:
5.02			
episode: 1655	score: 4.0	memory length: 381795	epsilon: 0.4420439200
104878	steps: 295	lr: 1.6000000000000003e-05	evaluation reward:
5.03			
episode: 1656	score: 11.0	memory length: 382324	epsilon: 0.440996500
0104812	steps: 529	lr: 1.6000000000000003e-05	evaluation reward:
5.07			
episode: 1657	score: 3.0	memory length: 382553	epsilon: 0.4405430800
104783	steps: 229	lr: 1.6000000000000003e-05	evaluation reward:
5.03			
episode: 1658	score: 5.0	memory length: 382898	epsilon: 0.4398599800
10474	steps: 345	lr: 1.6000000000000003e-05	evaluation reward: 5
.05			
episode: 1659	score: 10.0	memory length: 383290	epsilon: 0.439083820
0104691	steps: 392	lr: 1.6000000000000003e-05	evaluation reward:
5.11			
episode: 1660	score: 5.0	memory length: 383597	epsilon: 0.4384759600
1046524	steps: 307	lr: 1.6000000000000003e-05	evaluation reward:
5.1			
episode: 1661	score: 6.0	memory length: 383971	epsilon: 0.4377354400
1046055	steps: 374	lr: 1.6000000000000003e-05	evaluation reward:
5.13			

episode: 1662	score: 12.0	memory length: 384528	epsilon: 0.436632580
0104536	steps: 557	lr: 1.6000000000000003e-05	evaluation reward: 5.22
episode: 1663	score: 5.0	memory length: 384853	epsilon: 0.4359890800
104495	steps: 325	lr: 1.6000000000000003e-05	evaluation reward: 5.21
episode: 1664	score: 6.0	memory length: 385228	epsilon: 0.4352465800
104448	steps: 375	lr: 1.6000000000000003e-05	evaluation reward: 5.26
episode: 1665	score: 8.0	memory length: 385506	epsilon: 0.4346961400
104413	steps: 278	lr: 1.6000000000000003e-05	evaluation reward: 5.3
episode: 1666	score: 3.0	memory length: 385731	epsilon: 0.4342506400
104385	steps: 225	lr: 1.6000000000000003e-05	evaluation reward: 5.31
episode: 1667	score: 9.0	memory length: 386073	epsilon: 0.4335734800
104342	steps: 342	lr: 1.6000000000000003e-05	evaluation reward: 5.35
episode: 1668	score: 9.0	memory length: 386567	epsilon: 0.4325953600
1042803	steps: 494	lr: 1.6000000000000003e-05	evaluation reward: 5.38
episode: 1669	score: 6.0	memory length: 386940	epsilon: 0.4318568200
1042336	steps: 373	lr: 1.6000000000000003e-05	evaluation reward: 5.41
episode: 1670	score: 6.0	memory length: 387299	epsilon: 0.4311460000
1041886	steps: 359	lr: 1.6000000000000003e-05	evaluation reward: 5.41
episode: 1671	score: 10.0	memory length: 387852	epsilon: 0.430051060
01041193	steps: 553	lr: 1.6000000000000003e-05	evaluation reward : 5.47
episode: 1672	score: 12.0	memory length: 388325	epsilon: 0.429114520
010406	steps: 473	lr: 1.6000000000000003e-05	evaluation reward: 5.53
episode: 1673	score: 5.0	memory length: 388636	epsilon: 0.4284987400
104021	steps: 311	lr: 1.6000000000000003e-05	evaluation reward: 5.55
episode: 1674	score: 9.0	memory length: 389102	epsilon: 0.4275760600
103963	steps: 466	lr: 1.6000000000000003e-05	evaluation reward: 5.59
episode: 1675	score: 2.0	memory length: 389288	epsilon: 0.4272077800
1039394	steps: 186	lr: 1.6000000000000003e-05	evaluation reward: 5.59
episode: 1676	score: 8.0	memory length: 389761	epsilon: 0.4262712400
10388	steps: 473	lr: 1.6000000000000003e-05	evaluation reward: 5.62
episode: 1677	score: 9.0	memory length: 390232	epsilon: 0.4253386600
103821	steps: 471	lr: 1.6000000000000003e-05	evaluation reward: 5.68
episode: 1678	score: 7.0	memory length: 390593	epsilon: 0.4246238800
103776	steps: 361	lr: 1.6000000000000003e-05	evaluation reward: 5.69
episode: 1679	score: 8.0	memory length: 391008	epsilon: 0.4238021800
103724	steps: 415	lr: 1.6000000000000003e-05	evaluation reward:

5.69
episode: 1680 score: 6.0 memory length: 391348 epsilon: 0.4231289800
1036814 steps: 340 lr: 1.6000000000000003e-05 evaluation reward:
5.73
episode: 1681 score: 12.0 memory length: 391813 epsilon: 0.422208280
0103623 steps: 465 lr: 1.6000000000000003e-05 evaluation reward:
5.82
episode: 1682 score: 5.0 memory length: 392124 epsilon: 0.4215925000
103584 steps: 311 lr: 1.6000000000000003e-05 evaluation reward:
5.84
episode: 1683 score: 9.0 memory length: 392528 epsilon: 0.4207925800
1035336 steps: 404 lr: 1.6000000000000003e-05 evaluation reward:
5.88
episode: 1684 score: 7.0 memory length: 392934 epsilon: 0.4199887000
1034827 steps: 406 lr: 1.6000000000000003e-05 evaluation reward:
5.87
episode: 1685 score: 3.0 memory length: 393181 epsilon: 0.4194996400
103452 steps: 247 lr: 1.6000000000000003e-05 evaluation reward:
5.86
episode: 1686 score: 9.0 memory length: 393651 epsilon: 0.4185690400
103393 steps: 470 lr: 1.6000000000000003e-05 evaluation reward:
5.94
episode: 1687 score: 5.0 memory length: 393940 epsilon: 0.4179968200
1033567 steps: 289 lr: 1.6000000000000003e-05 evaluation reward:
5.96
episode: 1688 score: 8.0 memory length: 394416 epsilon: 0.4170543400
103297 steps: 476 lr: 1.6000000000000003e-05 evaluation reward:
5.93
episode: 1689 score: 5.0 memory length: 394724 epsilon: 0.4164445000
1032585 steps: 308 lr: 1.6000000000000003e-05 evaluation reward:
5.91
episode: 1690 score: 7.0 memory length: 395089 epsilon: 0.4157218000
103213 steps: 365 lr: 1.6000000000000003e-05 evaluation reward:
5.95
episode: 1691 score: 6.0 memory length: 395453 epsilon: 0.4150010800
103167 steps: 364 lr: 1.6000000000000003e-05 evaluation reward:
5.95
episode: 1692 score: 6.0 memory length: 395817 epsilon: 0.4142803600
1031215 steps: 364 lr: 1.6000000000000003e-05 evaluation reward:
5.93
episode: 1693 score: 8.0 memory length: 396217 epsilon: 0.4134883600
1030714 steps: 400 lr: 1.6000000000000003e-05 evaluation reward:
5.95
episode: 1694 score: 6.0 memory length: 396575 epsilon: 0.4127795200
1030266 steps: 358 lr: 1.6000000000000003e-05 evaluation reward:
5.95
episode: 1695 score: 6.0 memory length: 396915 epsilon: 0.4121063200
102984 steps: 340 lr: 1.6000000000000003e-05 evaluation reward:
5.87
episode: 1696 score: 3.0 memory length: 397127 epsilon: 0.4116865600
1029574 steps: 212 lr: 1.6000000000000003e-05 evaluation reward:
5.88
episode: 1697 score: 7.0 memory length: 397525 epsilon: 0.4108985200

1029076	steps: 398	lr: 1.6000000000000003e-05	evaluation reward:
5.88			
episode: 1698	score: 3.0	memory length: 397737	epsilon: 0.4104787600
102881	steps: 212	lr: 1.6000000000000003e-05	evaluation reward:
5.84			
episode: 1699	score: 5.0	memory length: 398059	epsilon: 0.4098412000
1028407	steps: 322	lr: 1.6000000000000003e-05	evaluation reward:
5.86			
episode: 1700	score: 6.0	memory length: 398434	epsilon: 0.4090987000
1027937	steps: 375	lr: 1.6000000000000003e-05	evaluation reward:
5.89			
episode: 1701	score: 3.0	memory length: 398644	epsilon: 0.4086829000
1027674	steps: 210	lr: 1.6000000000000003e-05	evaluation reward:
5.84			
episode: 1702	score: 8.0	memory length: 399056	epsilon: 0.4078671400
102716	steps: 412	lr: 1.6000000000000003e-05	evaluation reward:
5.85			
episode: 1703	score: 9.0	memory length: 399520	epsilon: 0.4069484200
1026577	steps: 464	lr: 1.6000000000000003e-05	evaluation reward:
5.89			
episode: 1704	score: 5.0	memory length: 399823	epsilon: 0.4063484800
1026197	steps: 303	lr: 1.6000000000000003e-05	evaluation reward:
5.89			
episode: 1705	score: 2.0	memory length: 400041	epsilon: 0.4059168400
1025924	steps: 218	lr: 6.400000000000001e-06	evaluation reward:
5.83			
episode: 1706	score: 9.0	memory length: 400528	epsilon: 0.4049525800
1025314	steps: 487	lr: 6.400000000000001e-06	evaluation reward:
5.89			
episode: 1707	score: 8.0	memory length: 400955	epsilon: 0.4041071200
102478	steps: 427	lr: 6.400000000000001e-06	evaluation reward: 5
.88			
episode: 1708	score: 4.0	memory length: 401234	epsilon: 0.4035547000
102443	steps: 279	lr: 6.400000000000001e-06	evaluation reward: 5
.86			
episode: 1709	score: 6.0	memory length: 401628	epsilon: 0.4027745800
1023936	steps: 394	lr: 6.400000000000001e-06	evaluation reward:
5.87			
episode: 1710	score: 5.0	memory length: 401971	epsilon: 0.4020954400
1023506	steps: 343	lr: 6.400000000000001e-06	evaluation reward:
5.89			
episode: 1711	score: 4.0	memory length: 402228	epsilon: 0.4015865800
1023184	steps: 257	lr: 6.400000000000001e-06	evaluation reward:
5.9			
episode: 1712	score: 8.0	memory length: 402666	epsilon: 0.4007193400
1022635	steps: 438	lr: 6.400000000000001e-06	evaluation reward:
5.91			
episode: 1713	score: 5.0	memory length: 402969	epsilon: 0.4001194000
1022256	steps: 303	lr: 6.400000000000001e-06	evaluation reward:
5.93			
episode: 1714	score: 4.0	memory length: 403266	epsilon: 0.3995313400
1021884	steps: 297	lr: 6.400000000000001e-06	evaluation reward:
5.91			

episode: 1715	score: 4.0	memory length: 403544	epsilon: 0.3989809000
1021536	steps: 278	lr: 6.400000000000001e-06	evaluation reward:
5.89			
episode: 1716	score: 5.0	memory length: 403868	epsilon: 0.3983393800
102113	steps: 324	lr: 6.400000000000001e-06	evaluation reward: 5
.91			
episode: 1717	score: 7.0	memory length: 404256	epsilon: 0.3975711400
1020644	steps: 388	lr: 6.400000000000001e-06	evaluation reward:
5.93			
episode: 1718	score: 5.0	memory length: 404599	epsilon: 0.3968920000
1020214	steps: 343	lr: 6.400000000000001e-06	evaluation reward:
5.93			
episode: 1719	score: 5.0	memory length: 404891	epsilon: 0.3963138400
101985	steps: 292	lr: 6.400000000000001e-06	evaluation reward: 5
.91			
episode: 1720	score: 9.0	memory length: 405362	epsilon: 0.3953812600
101926	steps: 471	lr: 6.400000000000001e-06	evaluation reward: 5
.94			
episode: 1721	score: 3.0	memory length: 405587	epsilon: 0.3949357600
1018976	steps: 225	lr: 6.400000000000001e-06	evaluation reward:
5.88			
episode: 1722	score: 2.0	memory length: 405768	epsilon: 0.3945773800
101875	steps: 181	lr: 6.400000000000001e-06	evaluation reward: 5
.83			
episode: 1723	score: 6.0	memory length: 406097	epsilon: 0.3939259600
101834	steps: 329	lr: 6.400000000000001e-06	evaluation reward: 5
.85			
episode: 1724	score: 9.0	memory length: 406584	epsilon: 0.3929617000
101773	steps: 487	lr: 6.400000000000001e-06	evaluation reward: 5
.91			
episode: 1725	score: 5.0	memory length: 406905	epsilon: 0.3923261200
1017325	steps: 321	lr: 6.400000000000001e-06	evaluation reward:
5.93			
episode: 1726	score: 6.0	memory length: 407227	epsilon: 0.3916885600
101692	steps: 322	lr: 6.400000000000001e-06	evaluation reward: 5
.95			
episode: 1727	score: 6.0	memory length: 407626	epsilon: 0.3908985400
101642	steps: 399	lr: 6.400000000000001e-06	evaluation reward: 5
.98			
episode: 1728	score: 8.0	memory length: 408047	epsilon: 0.3900649600
1015895	steps: 421	lr: 6.400000000000001e-06	evaluation reward:
5.99			
episode: 1729	score: 3.0	memory length: 408273	epsilon: 0.3896174800
101561	steps: 226	lr: 6.400000000000001e-06	evaluation reward: 5
.99			
episode: 1730	score: 6.0	memory length: 408646	epsilon: 0.3888789400
1015144	steps: 373	lr: 6.400000000000001e-06	evaluation reward:
6.0			
episode: 1731	score: 9.0	memory length: 409139	epsilon: 0.3879028000
1014527	steps: 493	lr: 6.400000000000001e-06	evaluation reward:
6.01			
episode: 1732	score: 8.0	memory length: 409548	epsilon: 0.3870929800
1014014	steps: 409	lr: 6.400000000000001e-06	evaluation reward:

6.04
episode: 1733 score: 5.0 memory length: 409858 epsilon: 0.3864791800
1013626 steps: 310 lr: 6.400000000000001e-06 evaluation reward:
6.05
episode: 1734 score: 3.0 memory length: 410088 epsilon: 0.3860237800
101334 steps: 230 lr: 6.400000000000001e-06 evaluation reward: 6
.05
episode: 1735 score: 6.0 memory length: 410437 epsilon: 0.3853327600
10129 steps: 349 lr: 6.400000000000001e-06 evaluation reward: 6.
06
episode: 1736 score: 7.0 memory length: 410791 epsilon: 0.3846318400
1012457 steps: 354 lr: 6.400000000000001e-06 evaluation reward:
6.04
episode: 1737 score: 7.0 memory length: 411186 epsilon: 0.3838497400
101196 steps: 395 lr: 6.400000000000001e-06 evaluation reward: 6
.07
episode: 1738 score: 3.0 memory length: 411432 epsilon: 0.3833626600
1011654 steps: 246 lr: 6.400000000000001e-06 evaluation reward:
6.04
episode: 1739 score: 5.0 memory length: 411743 epsilon: 0.3827468800
1011264 steps: 311 lr: 6.400000000000001e-06 evaluation reward:
6.03
episode: 1740 score: 10.0 memory length: 412211 epsilon: 0.381820240
0101068 steps: 468 lr: 6.400000000000001e-06 evaluation reward:
6.1
episode: 1741 score: 4.0 memory length: 412472 epsilon: 0.3813034600
101035 steps: 261 lr: 6.400000000000001e-06 evaluation reward: 6
.09
episode: 1742 score: 14.0 memory length: 413000 epsilon: 0.380258020
0100969 steps: 528 lr: 6.400000000000001e-06 evaluation reward:
6.19
episode: 1743 score: 3.0 memory length: 413212 epsilon: 0.3798382600
1009424 steps: 212 lr: 6.400000000000001e-06 evaluation reward:
6.14
episode: 1744 score: 11.0 memory length: 413625 epsilon: 0.379020520
01008907 steps: 413 lr: 6.400000000000001e-06 evaluation reward:
6.18
episode: 1745 score: 5.0 memory length: 413952 epsilon: 0.3783730600
1008497 steps: 327 lr: 6.400000000000001e-06 evaluation reward:
6.2
episode: 1746 score: 7.0 memory length: 414311 epsilon: 0.3776622400
100805 steps: 359 lr: 6.400000000000001e-06 evaluation reward: 6
.24
episode: 1747 score: 7.0 memory length: 414651 epsilon: 0.3769890400
100762 steps: 340 lr: 6.400000000000001e-06 evaluation reward: 6
.23
episode: 1748 score: 8.0 memory length: 415101 epsilon: 0.3760980400
100706 steps: 450 lr: 6.400000000000001e-06 evaluation reward: 6
.28
episode: 1749 score: 6.0 memory length: 415438 epsilon: 0.3754307800
1006636 steps: 337 lr: 6.400000000000001e-06 evaluation reward:
6.28
episode: 1750 score: 11.0 memory length: 415957 epsilon: 0.374403160

01005985	steps: 519	lr: 6.400000000000001e-06	evaluation reward:
6.33			
episode: 1751	score: 5.0	memory length: 416232	epsilon: 0.3738586600
100564	steps: 275	lr: 6.400000000000001e-06	evaluation reward: 6
.34			
episode: 1752	score: 6.0	memory length: 416611	epsilon: 0.3731082400
1005166	steps: 379	lr: 6.400000000000001e-06	evaluation reward:
6.37			
episode: 1753	score: 9.0	memory length: 417065	epsilon: 0.3722093200
10046	steps: 454	lr: 6.400000000000001e-06	evaluation reward: 6.
43			
episode: 1754	score: 6.0	memory length: 417412	epsilon: 0.3715222600
100416	steps: 347	lr: 6.400000000000001e-06	evaluation reward: 6
.44			
episode: 1755	score: 9.0	memory length: 417867	epsilon: 0.3706213600
100359	steps: 455	lr: 6.400000000000001e-06	evaluation reward: 6
.49			
episode: 1756	score: 6.0	memory length: 418238	epsilon: 0.3698867800
100313	steps: 371	lr: 6.400000000000001e-06	evaluation reward: 6
.44			
episode: 1757	score: 2.0	memory length: 418420	epsilon: 0.3695264200
10029	steps: 182	lr: 6.400000000000001e-06	evaluation reward: 6.
43			
episode: 1758	score: 3.0	memory length: 418633	epsilon: 0.3691046800
1002633	steps: 213	lr: 6.400000000000001e-06	evaluation reward:
6.41			
episode: 1759	score: 10.0	memory length: 419184	epsilon: 0.368013700
01001943	steps: 551	lr: 6.400000000000001e-06	evaluation reward:
6.41			
episode: 1760	score: 4.0	memory length: 419461	epsilon: 0.3674652400
1001596	steps: 277	lr: 6.400000000000001e-06	evaluation reward:
6.4			
episode: 1761	score: 7.0	memory length: 419842	epsilon: 0.3667108600
100112	steps: 381	lr: 6.400000000000001e-06	evaluation reward: 6
.41			
episode: 1762	score: 10.0	memory length: 420302	epsilon: 0.365800060
0100054	steps: 460	lr: 6.400000000000001e-06	evaluation reward:
6.39			
episode: 1763	score: 5.0	memory length: 420644	epsilon: 0.3651229000
1000114	steps: 342	lr: 6.400000000000001e-06	evaluation reward:
6.39			
episode: 1764	score: 12.0	memory length: 421248	epsilon: 0.363926980
0099936	steps: 604	lr: 6.400000000000001e-06	evaluation reward:
6.45			
episode: 1765	score: 3.0	memory length: 421478	epsilon: 0.3634715800
099907	steps: 230	lr: 6.400000000000001e-06	evaluation reward: 6
.4			
episode: 1766	score: 3.0	memory length: 421688	epsilon: 0.3630557800
0998806	steps: 210	lr: 6.400000000000001e-06	evaluation reward:
6.4			
episode: 1767	score: 3.0	memory length: 421898	epsilon: 0.3626399800
0998543	steps: 210	lr: 6.400000000000001e-06	evaluation reward:
6.34			

episode: 1768	score: 9.0	memory length: 422389	epsilon: 0.3616678000
099793	steps: 491	lr: 6.400000000000001e-06	evaluation reward: 6.34
episode: 1769	score: 10.0	memory length: 422871	epsilon: 0.360713440
00997324	steps: 482	lr: 6.400000000000001e-06	evaluation reward: 6.38
episode: 1770	score: 7.0	memory length: 423270	epsilon: 0.3599234200
0996824	steps: 399	lr: 6.400000000000001e-06	evaluation reward: 6.39
episode: 1771	score: 8.0	memory length: 423695	epsilon: 0.3590819200
099629	steps: 425	lr: 6.400000000000001e-06	evaluation reward: 6.37
episode: 1772	score: 3.0	memory length: 423907	epsilon: 0.3586621600
0996026	steps: 212	lr: 6.400000000000001e-06	evaluation reward: 6.28
episode: 1773	score: 11.0	memory length: 424460	epsilon: 0.357567220
00995333	steps: 553	lr: 6.400000000000001e-06	evaluation reward: 6.34
episode: 1774	score: 7.0	memory length: 424842	epsilon: 0.3568108600
0994855	steps: 382	lr: 6.400000000000001e-06	evaluation reward: 6.32
episode: 1775	score: 8.0	memory length: 425312	epsilon: 0.3558802600
0994266	steps: 470	lr: 6.400000000000001e-06	evaluation reward: 6.38
episode: 1776	score: 9.0	memory length: 425767	epsilon: 0.3549793600
0993696	steps: 455	lr: 6.400000000000001e-06	evaluation reward: 6.39
episode: 1777	score: 7.0	memory length: 426135	epsilon: 0.3542507200
0993235	steps: 368	lr: 6.400000000000001e-06	evaluation reward: 6.37
episode: 1778	score: 8.0	memory length: 426557	epsilon: 0.3534151600
0992707	steps: 422	lr: 6.400000000000001e-06	evaluation reward: 6.38
episode: 1779	score: 11.0	memory length: 426985	epsilon: 0.352567720
0099217	steps: 428	lr: 6.400000000000001e-06	evaluation reward: 6.41
episode: 1780	score: 7.0	memory length: 427388	epsilon: 0.3517697800
0991665	steps: 403	lr: 6.400000000000001e-06	evaluation reward: 6.42
episode: 1781	score: 8.0	memory length: 427828	epsilon: 0.3508985800
0991114	steps: 440	lr: 6.400000000000001e-06	evaluation reward: 6.38
episode: 1782	score: 5.0	memory length: 428116	epsilon: 0.3503283400
0990754	steps: 288	lr: 6.400000000000001e-06	evaluation reward: 6.38
episode: 1783	score: 6.0	memory length: 428457	epsilon: 0.3496531600
0990326	steps: 341	lr: 6.400000000000001e-06	evaluation reward: 6.35
episode: 1784	score: 4.0	memory length: 428719	epsilon: 0.3491344000
099	steps: 262	lr: 6.400000000000001e-06	evaluation reward: 6.32
episode: 1785	score: 6.0	memory length: 429072	epsilon: 0.3484354600
0989556	steps: 353	lr: 6.400000000000001e-06	evaluation reward: 6.35

episode: 1786	score: 9.0	memory length: 429567	epsilon: 0.3474553600
0988936	steps: 495	lr: 6.400000000000001e-06	evaluation reward:
6.35			
episode: 1787	score: 5.0	memory length: 429893	epsilon: 0.3468098800
098853	steps: 326	lr: 6.400000000000001e-06	evaluation reward: 6
.35			
episode: 1788	score: 5.0	memory length: 430258	epsilon: 0.3460871800
098807	steps: 365	lr: 6.400000000000001e-06	evaluation reward: 6
.32			
episode: 1789	score: 4.0	memory length: 430542	epsilon: 0.3455248600
0987714	steps: 284	lr: 6.400000000000001e-06	evaluation reward:
6.31			
episode: 1790	score: 3.0	memory length: 430755	epsilon: 0.3451031200
098745	steps: 213	lr: 6.400000000000001e-06	evaluation reward: 6
.27			
episode: 1791	score: 5.0	memory length: 431077	epsilon: 0.3444655600
0987044	steps: 322	lr: 6.400000000000001e-06	evaluation reward:
6.26			
episode: 1792	score: 11.0	memory length: 431602	epsilon: 0.343426060
00986386	steps: 525	lr: 6.400000000000001e-06	evaluation reward:
6.31			
episode: 1793	score: 11.0	memory length: 432089	epsilon: 0.342461800
00985776	steps: 487	lr: 6.400000000000001e-06	evaluation reward:
6.34			
episode: 1794	score: 9.0	memory length: 432577	epsilon: 0.3414955600
0985165	steps: 488	lr: 6.400000000000001e-06	evaluation reward:
6.37			
episode: 1795	score: 6.0	memory length: 432911	epsilon: 0.3408342400
0984747	steps: 334	lr: 6.400000000000001e-06	evaluation reward:
6.37			
episode: 1796	score: 7.0	memory length: 433314	epsilon: 0.3400363000
098424	steps: 403	lr: 6.400000000000001e-06	evaluation reward: 6
.41			
episode: 1797	score: 4.0	memory length: 433595	epsilon: 0.3394799200
098389	steps: 281	lr: 6.400000000000001e-06	evaluation reward: 6
.38			
episode: 1798	score: 6.0	memory length: 433924	epsilon: 0.3388285000
098348	steps: 329	lr: 6.400000000000001e-06	evaluation reward: 6
.41			
episode: 1799	score: 9.0	memory length: 434383	epsilon: 0.3379196800
09829	steps: 459	lr: 6.400000000000001e-06	evaluation reward: 6.
45			
episode: 1800	score: 3.0	memory length: 434599	epsilon: 0.3374920000
098263	steps: 216	lr: 6.400000000000001e-06	evaluation reward: 6
.42			
episode: 1801	score: 5.0	memory length: 434922	epsilon: 0.3368524600
098223	steps: 323	lr: 6.400000000000001e-06	evaluation reward: 6
.44			
episode: 1802	score: 4.0	memory length: 435201	epsilon: 0.3363000400
098188	steps: 279	lr: 6.400000000000001e-06	evaluation reward: 6
.4			
episode: 1803	score: 9.0	memory length: 435684	epsilon: 0.3353437000
0981273	steps: 483	lr: 6.400000000000001e-06	evaluation reward:

6.4
episode: 1804 score: 9.0 memory length: 436155 epsilon: 0.3344111200
0980683 steps: 471 lr: 6.400000000000001e-06 evaluation reward:
6.44
episode: 1805 score: 5.0 memory length: 436461 epsilon: 0.3338052400
09803 steps: 306 lr: 6.400000000000001e-06 evaluation reward: 6.
47
episode: 1806 score: 5.0 memory length: 436751 epsilon: 0.3332310400
0979936 steps: 290 lr: 6.400000000000001e-06 evaluation reward:
6.43
episode: 1807 score: 13.0 memory length: 437360 epsilon: 0.332025220
00979173 steps: 609 lr: 6.400000000000001e-06 evaluation reward:
6.48
episode: 1808 score: 5.0 memory length: 437671 epsilon: 0.3314094400
0978784 steps: 311 lr: 6.400000000000001e-06 evaluation reward:
6.49
episode: 1809 score: 8.0 memory length: 438123 epsilon: 0.3305144800
097822 steps: 452 lr: 6.400000000000001e-06 evaluation reward: 6
.51
episode: 1810 score: 12.0 memory length: 438570 epsilon: 0.329629420
0097766 steps: 447 lr: 6.400000000000001e-06 evaluation reward:
6.58
episode: 1811 score: 9.0 memory length: 438997 epsilon: 0.3287839600
097712 steps: 427 lr: 6.400000000000001e-06 evaluation reward: 6
.63
episode: 1812 score: 14.0 memory length: 439509 epsilon: 0.327770200
0097648 steps: 512 lr: 6.400000000000001e-06 evaluation reward:
6.69
episode: 1813 score: 5.0 memory length: 439798 epsilon: 0.3271979800
097612 steps: 289 lr: 6.400000000000001e-06 evaluation reward: 6
.69
episode: 1814 score: 5.0 memory length: 440112 epsilon: 0.3265762600
0975726 steps: 314 lr: 6.400000000000001e-06 evaluation reward:
6.7
episode: 1815 score: 9.0 memory length: 440585 epsilon: 0.3256397200
0975133 steps: 473 lr: 6.400000000000001e-06 evaluation reward:
6.75
episode: 1816 score: 8.0 memory length: 440998 epsilon: 0.3248219800
0974616 steps: 413 lr: 6.400000000000001e-06 evaluation reward:
6.78
episode: 1817 score: 6.0 memory length: 441336 epsilon: 0.3241527400
097419 steps: 338 lr: 6.400000000000001e-06 evaluation reward: 6
.77
episode: 1818 score: 5.0 memory length: 441659 epsilon: 0.3235132000
097379 steps: 323 lr: 6.400000000000001e-06 evaluation reward: 6
.77
episode: 1819 score: 7.0 memory length: 442065 epsilon: 0.3227093200
097328 steps: 406 lr: 6.400000000000001e-06 evaluation reward: 6
.79
episode: 1820 score: 4.0 memory length: 442343 epsilon: 0.3221588800
097293 steps: 278 lr: 6.400000000000001e-06 evaluation reward: 6
.74
episode: 1821 score: 2.0 memory length: 442527 epsilon: 0.3217945600

09727	steps: 184	lr: 6.400000000000001e-06	evaluation reward: 6.73
episode: 1822	score: 6.0	memory length: 442871	epsilon: 0.3211134400
097227	steps: 344	lr: 6.400000000000001e-06	evaluation reward: 6.77
episode: 1823	score: 7.0	memory length: 443277	epsilon: 0.3203095600
097176	steps: 406	lr: 6.400000000000001e-06	evaluation reward: 6.78
episode: 1824	score: 8.0	memory length: 443723	epsilon: 0.3194264800
09712	steps: 446	lr: 6.400000000000001e-06	evaluation reward: 6.77
episode: 1825	score: 13.0	memory length: 444194	epsilon: 0.318493900
0097061	steps: 471	lr: 6.400000000000001e-06	evaluation reward: 6.85
episode: 1826	score: 3.0	memory length: 444423	epsilon: 0.3180404800
0970325	steps: 229	lr: 6.400000000000001e-06	evaluation reward: 6.82
episode: 1827	score: 9.0	memory length: 444909	epsilon: 0.3170782000
0969716	steps: 486	lr: 6.400000000000001e-06	evaluation reward: 6.85
episode: 1828	score: 3.0	memory length: 445157	epsilon: 0.3165871600
0969406	steps: 248	lr: 6.400000000000001e-06	evaluation reward: 6.8
episode: 1829	score: 6.0	memory length: 445510	epsilon: 0.3158882200
0968964	steps: 353	lr: 6.400000000000001e-06	evaluation reward: 6.83
episode: 1830	score: 6.0	memory length: 445852	epsilon: 0.3152110600
0968535	steps: 342	lr: 6.400000000000001e-06	evaluation reward: 6.83
episode: 1831	score: 12.0	memory length: 446408	epsilon: 0.314110180
0096784	steps: 556	lr: 6.400000000000001e-06	evaluation reward: 6.86
episode: 1832	score: 5.0	memory length: 446680	epsilon: 0.3135716200
09675	steps: 272	lr: 6.400000000000001e-06	evaluation reward: 6.83
episode: 1833	score: 6.0	memory length: 447038	epsilon: 0.3128627800
096705	steps: 358	lr: 6.400000000000001e-06	evaluation reward: 6.84
episode: 1834	score: 4.0	memory length: 447319	epsilon: 0.3123064000
09667	steps: 281	lr: 6.400000000000001e-06	evaluation reward: 6.85
episode: 1835	score: 7.0	memory length: 447684	epsilon: 0.3115837000
096624	steps: 365	lr: 6.400000000000001e-06	evaluation reward: 6.86
episode: 1836	score: 4.0	memory length: 447966	epsilon: 0.3110253400
0965887	steps: 282	lr: 6.400000000000001e-06	evaluation reward: 6.83
episode: 1837	score: 8.0	memory length: 448393	epsilon: 0.3101798800
096535	steps: 427	lr: 6.400000000000001e-06	evaluation reward: 6.84
episode: 1838	score: 7.0	memory length: 448742	epsilon: 0.3094888600
0964915	steps: 349	lr: 6.400000000000001e-06	evaluation reward: 6.88

episode: 1839	score: 7.0	memory length: 449145	epsilon: 0.3086909200
096441	steps: 403	lr: 6.400000000000001e-06	evaluation reward: 6
.9			
episode: 1840	score: 7.0	memory length: 449527	epsilon: 0.3079345600
096393	steps: 382	lr: 6.400000000000001e-06	evaluation reward: 6
.87			
episode: 1841	score: 6.0	memory length: 449897	epsilon: 0.3072019600
096347	steps: 370	lr: 6.400000000000001e-06	evaluation reward: 6
.89			
episode: 1842	score: 4.0	memory length: 450178	epsilon: 0.3066455800
0963116	steps: 281	lr: 6.400000000000001e-06	evaluation reward:
6.79			
episode: 1843	score: 6.0	memory length: 450537	epsilon: 0.3059347600
0962666	steps: 359	lr: 6.400000000000001e-06	evaluation reward:
6.82			
episode: 1844	score: 11.0	memory length: 451114	epsilon: 0.304792300
00961943	steps: 577	lr: 6.400000000000001e-06	evaluation reward:
6.82			
episode: 1845	score: 4.0	memory length: 451392	epsilon: 0.3042418600
0961595	steps: 278	lr: 6.400000000000001e-06	evaluation reward:
6.81			
episode: 1846	score: 15.0	memory length: 451852	epsilon: 0.303331060
0096102	steps: 460	lr: 6.400000000000001e-06	evaluation reward:
6.89			
episode: 1847	score: 8.0	memory length: 452271	epsilon: 0.3025014400
0960494	steps: 419	lr: 6.400000000000001e-06	evaluation reward:
6.9			
episode: 1848	score: 4.0	memory length: 452514	epsilon: 0.3020203000
096019	steps: 243	lr: 6.400000000000001e-06	evaluation reward: 6
.86			
episode: 1849	score: 3.0	memory length: 452727	epsilon: 0.3015985600
095992	steps: 213	lr: 6.400000000000001e-06	evaluation reward: 6
.83			
episode: 1850	score: 8.0	memory length: 453147	epsilon: 0.3007669600
0959396	steps: 420	lr: 6.400000000000001e-06	evaluation reward:
6.8			
episode: 1851	score: 5.0	memory length: 453456	epsilon: 0.3001551400
095901	steps: 309	lr: 6.400000000000001e-06	evaluation reward: 6
.8			
episode: 1852	score: 5.0	memory length: 453781	epsilon: 0.2995116400
09586	steps: 325	lr: 6.400000000000001e-06	evaluation reward: 6.
79			
episode: 1853	score: 5.0	memory length: 454071	epsilon: 0.2989374400
095824	steps: 290	lr: 6.400000000000001e-06	evaluation reward: 6
.75			
episode: 1854	score: 3.0	memory length: 454302	epsilon: 0.2984800600
095795	steps: 231	lr: 6.400000000000001e-06	evaluation reward: 6
.72			
episode: 1855	score: 10.0	memory length: 454794	epsilon: 0.297505900
00957333	steps: 492	lr: 6.400000000000001e-06	evaluation reward:
6.73			
episode: 1856	score: 5.0	memory length: 455101	epsilon: 0.2968980400
095695	steps: 307	lr: 6.400000000000001e-06	evaluation reward: 6

.72
episode: 1857 score: 7.0 memory length: 455503 epsilon: 0.2961020800
0956445 steps: 402 lr: 6.400000000000001e-06 evaluation reward:
6.77
episode: 1858 score: 6.0 memory length: 455873 epsilon: 0.2953694800
095598 steps: 370 lr: 6.400000000000001e-06 evaluation reward: 6
.8
episode: 1859 score: 5.0 memory length: 456169 epsilon: 0.2947834000
095561 steps: 296 lr: 6.400000000000001e-06 evaluation reward: 6
.75
episode: 1860 score: 9.0 memory length: 456608 epsilon: 0.2939141800
095506 steps: 439 lr: 6.400000000000001e-06 evaluation reward: 6
.8
episode: 1861 score: 11.0 memory length: 457171 epsilon: 0.292799440
00954355 steps: 563 lr: 6.400000000000001e-06 evaluation reward:
6.84
episode: 1862 score: 8.0 memory length: 457591 epsilon: 0.2919678400
095383 steps: 420 lr: 6.400000000000001e-06 evaluation reward: 6
.82
episode: 1863 score: 4.0 memory length: 457865 epsilon: 0.2914253200
0953486 steps: 274 lr: 6.400000000000001e-06 evaluation reward:
6.81
episode: 1864 score: 7.0 memory length: 458270 epsilon: 0.2906234200
095298 steps: 405 lr: 6.400000000000001e-06 evaluation reward: 6
.76
episode: 1865 score: 4.0 memory length: 458544 epsilon: 0.2900809000
0952635 steps: 274 lr: 6.400000000000001e-06 evaluation reward:
6.77
episode: 1866 score: 3.0 memory length: 458773 epsilon: 0.2896274800
095235 steps: 229 lr: 6.400000000000001e-06 evaluation reward: 6
.77
episode: 1867 score: 5.0 memory length: 459076 epsilon: 0.2890275400
095197 steps: 303 lr: 6.400000000000001e-06 evaluation reward: 6
.79
episode: 1868 score: 6.0 memory length: 459449 epsilon: 0.2882890000
09515 steps: 373 lr: 6.400000000000001e-06 evaluation reward: 6.
76
episode: 1869 score: 8.0 memory length: 459910 epsilon: 0.2873762200
0950924 steps: 461 lr: 6.400000000000001e-06 evaluation reward:
6.74
episode: 1870 score: 3.0 memory length: 460143 epsilon: 0.2869148800
095063 steps: 233 lr: 6.400000000000001e-06 evaluation reward: 6
.7
episode: 1871 score: 9.0 memory length: 460611 epsilon: 0.2859882400
0950046 steps: 468 lr: 6.400000000000001e-06 evaluation reward:
6.71
episode: 1872 score: 8.0 memory length: 461040 epsilon: 0.2851388200
094951 steps: 429 lr: 6.400000000000001e-06 evaluation reward: 6
.76
episode: 1873 score: 7.0 memory length: 461449 epsilon: 0.2843290000
0948996 steps: 409 lr: 6.400000000000001e-06 evaluation reward:
6.72
episode: 1874 score: 7.0 memory length: 461857 epsilon: 0.2835211600

0948485	steps: 408	lr: 6.400000000000001e-06	evaluation reward:
6.72			
episode: 1875	score: 10.0	memory length: 462313	epsilon: 0.282618280
00947914	steps: 456	lr: 6.400000000000001e-06	evaluation reward:
6.74			
episode: 1876	score: 7.0	memory length: 462718	epsilon: 0.2818163800
0947407	steps: 405	lr: 6.400000000000001e-06	evaluation reward:
6.72			
episode: 1877	score: 5.0	memory length: 463020	epsilon: 0.2812184200
094703	steps: 302	lr: 6.400000000000001e-06	evaluation reward: 6
.7			
episode: 1878	score: 7.0	memory length: 463438	epsilon: 0.2803907800
0946505	steps: 418	lr: 6.400000000000001e-06	evaluation reward:
6.69			
episode: 1879	score: 8.0	memory length: 463836	epsilon: 0.2796027400
0946006	steps: 398	lr: 6.400000000000001e-06	evaluation reward:
6.66			
episode: 1880	score: 8.0	memory length: 464271	epsilon: 0.2787414400
094546	steps: 435	lr: 6.400000000000001e-06	evaluation reward: 6
.67			
episode: 1881	score: 3.0	memory length: 464518	epsilon: 0.2782523800
094515	steps: 247	lr: 6.400000000000001e-06	evaluation reward: 6
.62			
episode: 1882	score: 6.0	memory length: 464861	epsilon: 0.2775732400
094472	steps: 343	lr: 6.400000000000001e-06	evaluation reward: 6
.63			
episode: 1883	score: 3.0	memory length: 465107	epsilon: 0.2770861600
0944414	steps: 246	lr: 6.400000000000001e-06	evaluation reward:
6.6			
episode: 1884	score: 15.0	memory length: 465708	epsilon: 0.275896180
0094366	steps: 601	lr: 6.400000000000001e-06	evaluation reward:
6.71			
episode: 1885	score: 13.0	memory length: 466202	epsilon: 0.274918060
0094304	steps: 494	lr: 6.400000000000001e-06	evaluation reward:
6.78			
episode: 1886	score: 9.0	memory length: 466706	epsilon: 0.2739201400
094241	steps: 504	lr: 6.400000000000001e-06	evaluation reward: 6
.78			
episode: 1887	score: 15.0	memory length: 467256	epsilon: 0.272831140
0094172	steps: 550	lr: 6.400000000000001e-06	evaluation reward:
6.88			
episode: 1888	score: 7.0	memory length: 467639	epsilon: 0.2720728000
094124	steps: 383	lr: 6.400000000000001e-06	evaluation reward: 6
.9			
episode: 1889	score: 3.0	memory length: 467852	epsilon: 0.2716510600
0940975	steps: 213	lr: 6.400000000000001e-06	evaluation reward:
6.89			
episode: 1890	score: 5.0	memory length: 468177	epsilon: 0.2710075600
094057	steps: 325	lr: 6.400000000000001e-06	evaluation reward: 6
.91			
episode: 1891	score: 7.0	memory length: 468549	epsilon: 0.2702710000
09401	steps: 372	lr: 6.400000000000001e-06	evaluation reward: 6.
93			

episode: 1892	score: 9.0	memory length: 468952	epsilon: 0.2694730600
0939597	steps: 403	lr: 6.400000000000001e-06	evaluation reward:
6.91			
episode: 1893	score: 6.0	memory length: 469281	epsilon: 0.2688216400
0939185	steps: 329	lr: 6.400000000000001e-06	evaluation reward:
6.86			
episode: 1894	score: 4.0	memory length: 469559	epsilon: 0.2682712000
0938837	steps: 278	lr: 6.400000000000001e-06	evaluation reward:
6.81			
episode: 1895	score: 6.0	memory length: 469874	epsilon: 0.2676475000
093844	steps: 315	lr: 6.400000000000001e-06	evaluation reward: 6
.81			
episode: 1896	score: 8.0	memory length: 470298	epsilon: 0.2668079800
093791	steps: 424	lr: 6.400000000000001e-06	evaluation reward: 6
.82			
episode: 1897	score: 6.0	memory length: 470650	epsilon: 0.2661110200
093747	steps: 352	lr: 6.400000000000001e-06	evaluation reward: 6
.84			
episode: 1898	score: 5.0	memory length: 470923	epsilon: 0.2655704800
093713	steps: 273	lr: 6.400000000000001e-06	evaluation reward: 6
.83			
episode: 1899	score: 12.0	memory length: 471482	epsilon: 0.264463660
0093643	steps: 559	lr: 6.400000000000001e-06	evaluation reward:
6.86			
episode: 1900	score: 5.0	memory length: 471788	epsilon: 0.2638577800
0936044	steps: 306	lr: 6.400000000000001e-06	evaluation reward:
6.88			
episode: 1901	score: 7.0	memory length: 472152	epsilon: 0.2631370600
093559	steps: 364	lr: 6.400000000000001e-06	evaluation reward: 6
.9			
episode: 1902	score: 6.0	memory length: 472500	epsilon: 0.2624480200
093515	steps: 348	lr: 6.400000000000001e-06	evaluation reward: 6
.92			
episode: 1903	score: 7.0	memory length: 472852	epsilon: 0.2617510600
093471	steps: 352	lr: 6.400000000000001e-06	evaluation reward: 6
.9			
episode: 1904	score: 7.0	memory length: 473201	epsilon: 0.2610600400
0934274	steps: 349	lr: 6.400000000000001e-06	evaluation reward:
6.88			
episode: 1905	score: 7.0	memory length: 473586	epsilon: 0.2602977400
093379	steps: 385	lr: 6.400000000000001e-06	evaluation reward: 6
.9			
episode: 1906	score: 5.0	memory length: 473890	epsilon: 0.2596958200
093341	steps: 304	lr: 6.400000000000001e-06	evaluation reward: 6
.9			
episode: 1907	score: 13.0	memory length: 474358	epsilon: 0.258769180
00932825	steps: 468	lr: 6.400000000000001e-06	evaluation reward:
6.9			
episode: 1908	score: 4.0	memory length: 474652	epsilon: 0.2581870600
0932456	steps: 294	lr: 6.400000000000001e-06	evaluation reward:
6.89			
episode: 1909	score: 5.0	memory length: 474978	epsilon: 0.2575415800
093205	steps: 326	lr: 6.400000000000001e-06	evaluation reward: 6

```
.86
episode: 1910  score: 6.0  memory length: 475347  epsilon: 0.2568109600
0931586      steps: 369   lr: 6.400000000000001e-06  evaluation reward:
6.8
episode: 1911  score: 4.0  memory length: 475630  epsilon: 0.2562506200
093123      steps: 283   lr: 6.400000000000001e-06  evaluation reward: 6
.75
episode: 1912  score: 7.0  memory length: 476031  epsilon: 0.2554566400
093073      steps: 401   lr: 6.400000000000001e-06  evaluation reward: 6
.68
episode: 1913  score: 9.0  memory length: 476480  epsilon: 0.2545676200
0930167     steps: 449   lr: 6.400000000000001e-06  evaluation reward:
6.72
episode: 1914  score: 8.0  memory length: 476876  epsilon: 0.2537835400
092967     steps: 396   lr: 6.400000000000001e-06  evaluation reward: 6
.75
episode: 1915  score: 4.0  memory length: 477156  epsilon: 0.2532291400
092932     steps: 280   lr: 6.400000000000001e-06  evaluation reward: 6
.7
episode: 1916  score: 4.0  memory length: 477430  epsilon: 0.2526866200
0928976     steps: 274   lr: 6.400000000000001e-06  evaluation reward:
6.66
episode: 1917  score: 6.0  memory length: 477793  epsilon: 0.2519678800
092852     steps: 363   lr: 6.400000000000001e-06  evaluation reward: 6
.66
episode: 1918  score: 12.0 memory length: 478344  epsilon: 0.250876900
0092783     steps: 551   lr: 6.400000000000001e-06  evaluation reward:
6.73
episode: 1919  score: 9.0  memory length: 478801  epsilon: 0.2499720400
092726     steps: 457   lr: 6.400000000000001e-06  evaluation reward: 6
.75
episode: 1920  score: 5.0  memory length: 479108  epsilon: 0.2493641800
0926874     steps: 307   lr: 6.400000000000001e-06  evaluation reward:
6.76
episode: 1921  score: 5.0  memory length: 479411  epsilon: 0.2487642400
0926495     steps: 303   lr: 6.400000000000001e-06  evaluation reward:
6.79
episode: 1922  score: 11.0 memory length: 479842  epsilon: 0.247910860
00925955    steps: 431   lr: 6.400000000000001e-06  evaluation reward:
6.84
episode: 1923  score: 7.0  memory length: 480242  epsilon: 0.2471188600
0925454     steps: 400   lr: 6.400000000000001e-06  evaluation reward:
6.84
episode: 1924  score: 6.0  memory length: 480603  epsilon: 0.2464040800
0925001     steps: 361   lr: 6.400000000000001e-06  evaluation reward:
6.82
episode: 1925  score: 4.0  memory length: 480901  epsilon: 0.2458140400
0924628     steps: 298   lr: 6.400000000000001e-06  evaluation reward:
6.73
episode: 1926  score: 7.0  memory length: 481306  epsilon: 0.2450121400
092412     steps: 405   lr: 6.400000000000001e-06  evaluation reward: 6
.77
episode: 1927  score: 9.0  memory length: 481755  epsilon: 0.2441231200
```


0923558	steps: 449	lr: 6.400000000000001e-06	evaluation reward:
6.77			
episode: 1928	score: 5.0	memory length: 482077	epsilon: 0.2434855600
0923155	steps: 322	lr: 6.400000000000001e-06	evaluation reward:
6.79			
episode: 1929	score: 7.0	memory length: 482470	epsilon: 0.2427074200
0922663	steps: 393	lr: 6.400000000000001e-06	evaluation reward:
6.8			
episode: 1930	score: 9.0	memory length: 482919	epsilon: 0.2418184000
09221	steps: 449	lr: 6.400000000000001e-06	evaluation reward: 6.
83			
episode: 1931	score: 10.0	memory length: 483438	epsilon: 0.240790780
0092145	steps: 519	lr: 6.400000000000001e-06	evaluation reward:
6.81			
episode: 1932	score: 4.0	memory length: 483680	epsilon: 0.2403116200
0921147	steps: 242	lr: 6.400000000000001e-06	evaluation reward:
6.8			
episode: 1933	score: 9.0	memory length: 484143	epsilon: 0.2393948800
0920567	steps: 463	lr: 6.400000000000001e-06	evaluation reward:
6.83			
episode: 1934	score: 6.0	memory length: 484516	epsilon: 0.2386563400
09201	steps: 373	lr: 6.400000000000001e-06	evaluation reward: 6.
85			
episode: 1935	score: 3.0	memory length: 484764	epsilon: 0.2381653000
091979	steps: 248	lr: 6.400000000000001e-06	evaluation reward: 6
.81			
episode: 1936	score: 5.0	memory length: 485056	epsilon: 0.2375871400
0919423	steps: 292	lr: 6.400000000000001e-06	evaluation reward:
6.82			
episode: 1937	score: 8.0	memory length: 485514	epsilon: 0.2366803000
091885	steps: 458	lr: 6.400000000000001e-06	evaluation reward: 6
.82			
episode: 1938	score: 14.0	memory length: 486070	epsilon: 0.235579420
00918153	steps: 556	lr: 6.400000000000001e-06	evaluation reward:
6.89			
episode: 1939	score: 6.0	memory length: 486423	epsilon: 0.2348804800
091771	steps: 353	lr: 6.400000000000001e-06	evaluation reward: 6
.88			
episode: 1940	score: 8.0	memory length: 486882	epsilon: 0.2339716600
0917136	steps: 459	lr: 6.400000000000001e-06	evaluation reward:
6.89			
episode: 1941	score: 4.0	memory length: 487181	epsilon: 0.2333796400
091676	steps: 299	lr: 6.400000000000001e-06	evaluation reward: 6
.87			
episode: 1942	score: 8.0	memory length: 487616	epsilon: 0.2325183400
0916216	steps: 435	lr: 6.400000000000001e-06	evaluation reward:
6.91			
episode: 1943	score: 9.0	memory length: 488072	epsilon: 0.2316154600
0915645	steps: 456	lr: 6.400000000000001e-06	evaluation reward:
6.94			
episode: 1944	score: 10.0	memory length: 488442	epsilon: 0.230882860
0091518	steps: 370	lr: 6.400000000000001e-06	evaluation reward:
6.93			

episode: 1945	score: 5.0	memory length: 488714	epsilon: 0.2303443000
091484	steps: 272	lr: 6.400000000000001e-06	evaluation reward: 6
.94			
episode: 1946	score: 4.0	memory length: 488957	epsilon: 0.2298631600
0914536	steps: 243	lr: 6.400000000000001e-06	evaluation reward:
6.83			
episode: 1947	score: 6.0	memory length: 489329	epsilon: 0.2291266000
091407	steps: 372	lr: 6.400000000000001e-06	evaluation reward: 6
.81			
episode: 1948	score: 12.0	memory length: 489808	epsilon: 0.228178180
0091347	steps: 479	lr: 6.400000000000001e-06	evaluation reward:
6.89			
episode: 1949	score: 7.0	memory length: 490196	epsilon: 0.2274099400
0912984	steps: 388	lr: 6.400000000000001e-06	evaluation reward:
6.93			
episode: 1950	score: 14.0	memory length: 490789	epsilon: 0.226235800
0091224	steps: 593	lr: 6.400000000000001e-06	evaluation reward:
6.99			
episode: 1951	score: 9.0	memory length: 491218	epsilon: 0.2253863800
0911704	steps: 429	lr: 6.400000000000001e-06	evaluation reward:
7.03			
episode: 1952	score: 13.0	memory length: 491818	epsilon: 0.224198380
00910952	steps: 600	lr: 6.400000000000001e-06	evaluation reward:
7.11			
episode: 1953	score: 11.0	memory length: 492351	epsilon: 0.223143040
00910284	steps: 533	lr: 6.400000000000001e-06	evaluation reward:
7.17			
episode: 1954	score: 5.0	memory length: 492636	epsilon: 0.2225787400
0909927	steps: 285	lr: 6.400000000000001e-06	evaluation reward:
7.19			
episode: 1955	score: 9.0	memory length: 493105	epsilon: 0.2216501200
090934	steps: 469	lr: 6.400000000000001e-06	evaluation reward: 7
.18			
episode: 1956	score: 7.0	memory length: 493494	epsilon: 0.2208799000
0908853	steps: 389	lr: 6.400000000000001e-06	evaluation reward:
7.2			
episode: 1957	score: 9.0	memory length: 493931	epsilon: 0.2200146400
0908305	steps: 437	lr: 6.400000000000001e-06	evaluation reward:
7.22			
episode: 1958	score: 4.0	memory length: 494208	epsilon: 0.2194661800
0907958	steps: 277	lr: 6.400000000000001e-06	evaluation reward:
7.2			
episode: 1959	score: 6.0	memory length: 494583	epsilon: 0.2187236800
0907488	steps: 375	lr: 6.400000000000001e-06	evaluation reward:
7.21			
episode: 1960	score: 8.0	memory length: 495041	epsilon: 0.2178168400
0906915	steps: 458	lr: 6.400000000000001e-06	evaluation reward:
7.2			
episode: 1961	score: 4.0	memory length: 495319	epsilon: 0.2172664000
0906566	steps: 278	lr: 6.400000000000001e-06	evaluation reward:
7.13			
episode: 1962	score: 11.0	memory length: 495856	epsilon: 0.216203140
00905894	steps: 537	lr: 6.400000000000001e-06	evaluation reward:

7.16
episode: 1963 score: 5.0 memory length: 496144 epsilon: 0.2156329000
0905533 steps: 288 lr: 6.400000000000001e-06 evaluation reward:
7.17
episode: 1964 score: 8.0 memory length: 496589 epsilon: 0.2147518000
0904975 steps: 445 lr: 6.400000000000001e-06 evaluation reward:
7.18
episode: 1965 score: 7.0 memory length: 496992 epsilon: 0.2139538600
090447 steps: 403 lr: 6.400000000000001e-06 evaluation reward: 7
.21
episode: 1966 score: 4.0 memory length: 497269 epsilon: 0.2134054000
0904124 steps: 277 lr: 6.400000000000001e-06 evaluation reward:
7.22
episode: 1967 score: 8.0 memory length: 497685 epsilon: 0.2125817200
0903602 steps: 416 lr: 6.400000000000001e-06 evaluation reward:
7.25
episode: 1968 score: 9.0 memory length: 497997 epsilon: 0.2119639600
0903212 steps: 312 lr: 6.400000000000001e-06 evaluation reward:
7.28
episode: 1969 score: 15.0 memory length: 498593 epsilon: 0.210783880
00902465 steps: 596 lr: 6.400000000000001e-06 evaluation reward:
7.35
episode: 1970 score: 5.0 memory length: 498877 epsilon: 0.2102215600
090211 steps: 284 lr: 6.400000000000001e-06 evaluation reward: 7
.37
episode: 1971 score: 4.0 memory length: 499120 epsilon: 0.2097404200
0901805 steps: 243 lr: 6.400000000000001e-06 evaluation reward:
7.32
episode: 1972 score: 7.0 memory length: 499506 epsilon: 0.2089761400
090132 steps: 386 lr: 6.400000000000001e-06 evaluation reward: 7
.31
episode: 1973 score: 5.0 memory length: 499815 epsilon: 0.2083643200
0900934 steps: 309 lr: 6.400000000000001e-06 evaluation reward:
7.29
episode: 1974 score: 17.0 memory length: 500364 epsilon: 0.207277300
00900246 steps: 549 lr: 2.560000000000001e-06 evaluation reward:
7.39
episode: 1975 score: 8.0 memory length: 500809 epsilon: 0.2063962000
089969 steps: 445 lr: 2.560000000000001e-06 evaluation reward: 7
.37
episode: 1976 score: 9.0 memory length: 501289 epsilon: 0.2054458000
0899088 steps: 480 lr: 2.560000000000001e-06 evaluation reward:
7.39
episode: 1977 score: 12.0 memory length: 501769 epsilon: 0.204495400
00898486 steps: 480 lr: 2.560000000000001e-06 evaluation reward:
7.46
episode: 1978 score: 5.0 memory length: 502058 epsilon: 0.2039231800
0898124 steps: 289 lr: 2.560000000000001e-06 evaluation reward:
7.44
episode: 1979 score: 8.0 memory length: 502491 epsilon: 0.2030658400
0897582 steps: 433 lr: 2.560000000000001e-06 evaluation reward:
7.44
episode: 1980 score: 4.0 memory length: 502750 epsilon: 0.2025530200

0897257	steps: 259	lr: 2.560000000000001e-06	evaluation reward:
7.4			
episode: 1981	score: 11.0	memory length: 503170	epsilon: 0.201721420
0089673	steps: 420	lr: 2.560000000000001e-06	evaluation reward:
7.48			
episode: 1982	score: 5.0	memory length: 503517	epsilon: 0.2010343600
0896296	steps: 347	lr: 2.560000000000001e-06	evaluation reward:
7.47			
episode: 1983	score: 7.0	memory length: 503906	epsilon: 0.2002641400
089581	steps: 389	lr: 2.560000000000001e-06	evaluation reward: 7
.51			
episode: 1984	score: 9.0	memory length: 504409	epsilon: 0.1992682000
089518	steps: 503	lr: 2.560000000000001e-06	evaluation reward: 7
.45			
episode: 1985	score: 6.0	memory length: 504752	epsilon: 0.1985890600
089475	steps: 343	lr: 2.560000000000001e-06	evaluation reward: 7
.38			
episode: 1986	score: 9.0	memory length: 505181	epsilon: 0.1977396400
0894212	steps: 429	lr: 2.560000000000001e-06	evaluation reward:
7.38			
episode: 1987	score: 9.0	memory length: 505657	epsilon: 0.1967971600
0893616	steps: 476	lr: 2.560000000000001e-06	evaluation reward:
7.32			
episode: 1988	score: 9.0	memory length: 505985	epsilon: 0.1961477200
0893205	steps: 328	lr: 2.560000000000001e-06	evaluation reward:
7.34			
episode: 1989	score: 8.0	memory length: 506343	epsilon: 0.1954388800
0892756	steps: 358	lr: 2.560000000000001e-06	evaluation reward:
7.39			
episode: 1990	score: 7.0	memory length: 506718	epsilon: 0.1946963800
0892286	steps: 375	lr: 2.560000000000001e-06	evaluation reward:
7.41			
episode: 1991	score: 10.0	memory length: 507231	epsilon: 0.193680640
00891644	steps: 513	lr: 2.560000000000001e-06	evaluation reward:
7.44			
episode: 1992	score: 8.0	memory length: 507664	epsilon: 0.1928233000
08911	steps: 433	lr: 2.560000000000001e-06	evaluation reward: 7.
43			
episode: 1993	score: 10.0	memory length: 508130	epsilon: 0.191900620
00890518	steps: 466	lr: 2.560000000000001e-06	evaluation reward:
7.47			
episode: 1994	score: 9.0	memory length: 508606	epsilon: 0.1909581400
088992	steps: 476	lr: 2.560000000000001e-06	evaluation reward: 7
.52			
episode: 1995	score: 5.0	memory length: 508909	epsilon: 0.1903582000
0889542	steps: 303	lr: 2.560000000000001e-06	evaluation reward:
7.51			
episode: 1996	score: 4.0	memory length: 509185	epsilon: 0.1898117200
0889196	steps: 276	lr: 2.560000000000001e-06	evaluation reward:
7.47			
episode: 1997	score: 9.0	memory length: 509635	epsilon: 0.1889207200
0888632	steps: 450	lr: 2.560000000000001e-06	evaluation reward:
7.5			

episode: 1998	score: 10.0	memory length: 510138	epsilon: 0.187924780
00888002	steps: 503	lr: 2.560000000000001e-06	evaluation reward: 7.55
episode: 1999	score: 10.0	memory length: 510650	epsilon: 0.186911020
0088736	steps: 512	lr: 2.560000000000001e-06	evaluation reward: 7.53
episode: 2000	score: 7.0	memory length: 511059	epsilon: 0.1861012000
0886848	steps: 409	lr: 2.560000000000001e-06	evaluation reward: 7.55
episode: 2001	score: 9.0	memory length: 511507	epsilon: 0.1852141600
0886287	steps: 448	lr: 2.560000000000001e-06	evaluation reward: 7.57
episode: 2002	score: 9.0	memory length: 511980	epsilon: 0.1842776200
0885695	steps: 473	lr: 2.560000000000001e-06	evaluation reward: 7.6
episode: 2003	score: 7.0	memory length: 512353	epsilon: 0.1835390800
0885227	steps: 373	lr: 2.560000000000001e-06	evaluation reward: 7.6
episode: 2004	score: 6.0	memory length: 512696	epsilon: 0.1828599400
0884798	steps: 343	lr: 2.560000000000001e-06	evaluation reward: 7.59
episode: 2005	score: 7.0	memory length: 513120	epsilon: 0.1820204200
0884266	steps: 424	lr: 2.560000000000001e-06	evaluation reward: 7.59
episode: 2006	score: 6.0	memory length: 513460	epsilon: 0.1813472200
088384	steps: 340	lr: 2.560000000000001e-06	evaluation reward: 7.6
episode: 2007	score: 10.0	memory length: 513988	epsilon: 0.180301780
0088318	steps: 528	lr: 2.560000000000001e-06	evaluation reward: 7.57
episode: 2008	score: 8.0	memory length: 514443	epsilon: 0.1794008800
088261	steps: 455	lr: 2.560000000000001e-06	evaluation reward: 7.61
episode: 2009	score: 11.0	memory length: 515000	epsilon: 0.178298020
0088191	steps: 557	lr: 2.560000000000001e-06	evaluation reward: 7.67
episode: 2010	score: 11.0	memory length: 515553	epsilon: 0.177203080
00881219	steps: 553	lr: 2.560000000000001e-06	evaluation reward: 7.72
episode: 2011	score: 8.0	memory length: 515941	epsilon: 0.1764348400
0880733	steps: 388	lr: 2.560000000000001e-06	evaluation reward: 7.76
episode: 2012	score: 6.0	memory length: 516275	epsilon: 0.1757735200
0880314	steps: 334	lr: 2.560000000000001e-06	evaluation reward: 7.75
episode: 2013	score: 9.0	memory length: 516749	epsilon: 0.1748350000
087972	steps: 474	lr: 2.560000000000001e-06	evaluation reward: 7.75
episode: 2014	score: 7.0	memory length: 517111	epsilon: 0.1741182400
0879267	steps: 362	lr: 2.560000000000001e-06	evaluation reward: 7.74
episode: 2015	score: 9.0	memory length: 517553	epsilon: 0.1732430800
0878713	steps: 442	lr: 2.560000000000001e-06	evaluation reward:

7.79
episode: 2016 score: 5.0 memory length: 517875 epsilon: 0.1726055200
087831 steps: 322 lr: 2.560000000000001e-06 evaluation reward: 7
.8
episode: 2017 score: 6.0 memory length: 518225 epsilon: 0.1719125200
087787 steps: 350 lr: 2.560000000000001e-06 evaluation reward: 7
.8
episode: 2018 score: 4.0 memory length: 518500 epsilon: 0.1713680200
0877527 steps: 275 lr: 2.560000000000001e-06 evaluation reward:
7.72
episode: 2019 score: 8.0 memory length: 518912 epsilon: 0.1705522600
087701 steps: 412 lr: 2.560000000000001e-06 evaluation reward: 7
.71
episode: 2020 score: 7.0 memory length: 519270 epsilon: 0.1698434200
0876562 steps: 358 lr: 2.560000000000001e-06 evaluation reward:
7.73
episode: 2021 score: 10.0 memory length: 519765 epsilon: 0.168863320
00875942 steps: 495 lr: 2.560000000000001e-06 evaluation reward:
7.78
episode: 2022 score: 9.0 memory length: 520245 epsilon: 0.1679129200
087534 steps: 480 lr: 2.560000000000001e-06 evaluation reward: 7
.76
episode: 2023 score: 5.0 memory length: 520551 epsilon: 0.1673070400
0874957 steps: 306 lr: 2.560000000000001e-06 evaluation reward:
7.74
episode: 2024 score: 5.0 memory length: 520880 epsilon: 0.1666556200
0874545 steps: 329 lr: 2.560000000000001e-06 evaluation reward:
7.73
episode: 2025 score: 7.0 memory length: 521280 epsilon: 0.1658636200
0874044 steps: 400 lr: 2.560000000000001e-06 evaluation reward:
7.76
episode: 2026 score: 7.0 memory length: 521708 epsilon: 0.1650161800
0873508 steps: 428 lr: 2.560000000000001e-06 evaluation reward:
7.76
episode: 2027 score: 8.0 memory length: 522133 epsilon: 0.1641746800
0872976 steps: 425 lr: 2.560000000000001e-06 evaluation reward:
7.75
episode: 2028 score: 8.0 memory length: 522559 epsilon: 0.1633312000
0872442 steps: 426 lr: 2.560000000000001e-06 evaluation reward:
7.78
episode: 2029 score: 11.0 memory length: 523134 epsilon: 0.162192700
00871722 steps: 575 lr: 2.560000000000001e-06 evaluation reward:
7.82
episode: 2030 score: 6.0 memory length: 523488 epsilon: 0.1614917800
0871278 steps: 354 lr: 2.560000000000001e-06 evaluation reward:
7.79
episode: 2031 score: 7.0 memory length: 523892 epsilon: 0.1606918600
0870772 steps: 404 lr: 2.560000000000001e-06 evaluation reward:
7.76
episode: 2032 score: 7.0 memory length: 524284 epsilon: 0.1599157000
087028 steps: 392 lr: 2.560000000000001e-06 evaluation reward: 7
.79
episode: 2033 score: 4.0 memory length: 524584 epsilon: 0.1593217000

0869905	steps: 300	lr: 2.5600000000000001e-06	evaluation reward:
7.74			
episode: 2034	score: 7.0	memory length: 524952	epsilon: 0.1585930600
0869444	steps: 368	lr: 2.5600000000000001e-06	evaluation reward:
7.75			
episode: 2035	score: 9.0	memory length: 525446	epsilon: 0.1576149400
0868825	steps: 494	lr: 2.5600000000000001e-06	evaluation reward:
7.81			
episode: 2036	score: 7.0	memory length: 525816	epsilon: 0.1568823400
0868362	steps: 370	lr: 2.5600000000000001e-06	evaluation reward:
7.83			
episode: 2037	score: 7.0	memory length: 526238	epsilon: 0.1560467800
0867833	steps: 422	lr: 2.5600000000000001e-06	evaluation reward:
7.82			
episode: 2038	score: 9.0	memory length: 526653	epsilon: 0.1552250800
0867313	steps: 415	lr: 2.5600000000000001e-06	evaluation reward:
7.77			
episode: 2039	score: 7.0	memory length: 527038	epsilon: 0.1544627800
086683	steps: 385	lr: 2.5600000000000001e-06	evaluation reward: 7
.78			
episode: 2040	score: 10.0	memory length: 527609	epsilon: 0.153332200
00866116	steps: 571	lr: 2.5600000000000001e-06	evaluation reward:
7.8			
episode: 2041	score: 9.0	memory length: 528103	epsilon: 0.1523540800
0865497	steps: 494	lr: 2.5600000000000001e-06	evaluation reward:
7.85			
episode: 2042	score: 8.0	memory length: 528518	epsilon: 0.1515323800
0864977	steps: 415	lr: 2.5600000000000001e-06	evaluation reward:
7.85			
episode: 2043	score: 4.0	memory length: 528795	epsilon: 0.1509839200
086463	steps: 277	lr: 2.5600000000000001e-06	evaluation reward: 7
.8			
episode: 2044	score: 8.0	memory length: 529214	epsilon: 0.1501543000
0864105	steps: 419	lr: 2.5600000000000001e-06	evaluation reward:
7.78			
episode: 2045	score: 3.0	memory length: 529461	epsilon: 0.1496652400
0863796	steps: 247	lr: 2.5600000000000001e-06	evaluation reward:
7.76			
episode: 2046	score: 11.0	memory length: 529985	epsilon: 0.148627720
0086314	steps: 524	lr: 2.5600000000000001e-06	evaluation reward:
7.83			
episode: 2047	score: 3.0	memory length: 530232	epsilon: 0.1481386600
086283	steps: 247	lr: 2.5600000000000001e-06	evaluation reward: 7
.8			
episode: 2048	score: 16.0	memory length: 530852	epsilon: 0.146911060
00862053	steps: 620	lr: 2.5600000000000001e-06	evaluation reward:
7.84			
episode: 2049	score: 6.0	memory length: 531168	epsilon: 0.1462853800
0861657	steps: 316	lr: 2.5600000000000001e-06	evaluation reward:
7.83			
episode: 2050	score: 10.0	memory length: 531544	epsilon: 0.145540900
00861186	steps: 376	lr: 2.5600000000000001e-06	evaluation reward:
7.79			

episode: 2051	score: 9.0	memory length: 532034	epsilon: 0.1445707000
0860572	steps: 490	lr: 2.560000000000001e-06	evaluation reward:
7.79			
episode: 2052	score: 5.0	memory length: 532341	epsilon: 0.1439628400
0860188	steps: 307	lr: 2.560000000000001e-06	evaluation reward:
7.71			
episode: 2053	score: 12.0	memory length: 532921	epsilon: 0.142814440
0085946	steps: 580	lr: 2.560000000000001e-06	evaluation reward:
7.72			
episode: 2054	score: 4.0	memory length: 533215	epsilon: 0.1422323200
0859093	steps: 294	lr: 2.560000000000001e-06	evaluation reward:
7.71			
episode: 2055	score: 11.0	memory length: 533766	epsilon: 0.141141340
00858403	steps: 551	lr: 2.560000000000001e-06	evaluation reward:
7.73			
episode: 2056	score: 12.0	memory length: 534244	epsilon: 0.140194900
00857804	steps: 478	lr: 2.560000000000001e-06	evaluation reward:
7.78			
episode: 2057	score: 11.0	memory length: 534766	epsilon: 0.139161340
0085715	steps: 522	lr: 2.560000000000001e-06	evaluation reward:
7.8			
episode: 2058	score: 4.0	memory length: 535039	epsilon: 0.1386208000
0856808	steps: 273	lr: 2.560000000000001e-06	evaluation reward:
7.8			
episode: 2059	score: 4.0	memory length: 535302	epsilon: 0.1381000600
0856478	steps: 263	lr: 2.560000000000001e-06	evaluation reward:
7.78			
episode: 2060	score: 7.0	memory length: 535704	epsilon: 0.1373041000
0855975	steps: 402	lr: 2.560000000000001e-06	evaluation reward:
7.77			
episode: 2061	score: 7.0	memory length: 536082	epsilon: 0.1365556600
08555	steps: 378	lr: 2.560000000000001e-06	evaluation reward: 7.
8			
episode: 2062	score: 6.0	memory length: 536438	epsilon: 0.1358507800
0855055	steps: 356	lr: 2.560000000000001e-06	evaluation reward:
7.75			
episode: 2063	score: 9.0	memory length: 536892	epsilon: 0.1349518600
0854487	steps: 454	lr: 2.560000000000001e-06	evaluation reward:
7.79			
episode: 2064	score: 6.0	memory length: 537254	epsilon: 0.1342351000
0854033	steps: 362	lr: 2.560000000000001e-06	evaluation reward:
7.77			
episode: 2065	score: 9.0	memory length: 537693	epsilon: 0.1333658800
0853483	steps: 439	lr: 2.560000000000001e-06	evaluation reward:
7.79			
episode: 2066	score: 9.0	memory length: 538105	epsilon: 0.1325501200
0852967	steps: 412	lr: 2.560000000000001e-06	evaluation reward:
7.84			
episode: 2067	score: 12.0	memory length: 538767	epsilon: 0.131239360
00852138	steps: 662	lr: 2.560000000000001e-06	evaluation reward:
7.88			
episode: 2068	score: 9.0	memory length: 539227	epsilon: 0.1303285600
0851561	steps: 460	lr: 2.560000000000001e-06	evaluation reward:


```
7.88
episode: 2069    score: 5.0    memory length: 539533    epsilon: 0.1297226800
0851178      steps: 306    lr: 2.5600000000000001e-06    evaluation reward:
7.78
episode: 2070    score: 8.0    memory length: 539977    epsilon: 0.1288435600
0850622      steps: 444    lr: 2.5600000000000001e-06    evaluation reward:
7.81
episode: 2071    score: 9.0    memory length: 540443    epsilon: 0.1279208800
0850038      steps: 466    lr: 2.5600000000000001e-06    evaluation reward:
7.86
episode: 2072    score: 7.0    memory length: 540828    epsilon: 0.1271585800
0849556      steps: 385    lr: 2.5600000000000001e-06    evaluation reward:
7.86
episode: 2073    score: 6.0    memory length: 541203    epsilon: 0.1264160800
0849086      steps: 375    lr: 2.5600000000000001e-06    evaluation reward:
7.87
episode: 2074    score: 8.0    memory length: 541657    epsilon: 0.1255171600
0848517      steps: 454    lr: 2.5600000000000001e-06    evaluation reward:
7.78
episode: 2075    score: 8.0    memory length: 542091    epsilon: 0.1246578400
0848214      steps: 434    lr: 2.5600000000000001e-06    evaluation reward:
7.78
episode: 2076    score: 11.0   memory length: 542664    epsilon: 0.123523300
00848291     steps: 573    lr: 2.5600000000000001e-06    evaluation reward:
7.8
episode: 2077    score: 7.0    memory length: 543019    epsilon: 0.1228204000
0848339      steps: 355    lr: 2.5600000000000001e-06    evaluation reward:
7.75
episode: 2078    score: 8.0    memory length: 543514    epsilon: 0.1218403000
0848406      steps: 495    lr: 2.5600000000000001e-06    evaluation reward:
7.78
episode: 2079    score: 7.0    memory length: 543896    epsilon: 0.1210839400
0848458      steps: 382    lr: 2.5600000000000001e-06    evaluation reward:
7.77
episode: 2080    score: 10.0   memory length: 544399    epsilon: 0.120088000
00848525     steps: 503    lr: 2.5600000000000001e-06    evaluation reward:
7.83
episode: 2081    score: 7.0    memory length: 544759    epsilon: 0.1193752000
0848574      steps: 360    lr: 2.5600000000000001e-06    evaluation reward:
7.79
episode: 2082    score: 6.0    memory length: 545132    epsilon: 0.1186366600
0848624      steps: 373    lr: 2.5600000000000001e-06    evaluation reward:
7.8
episode: 2083    score: 8.0    memory length: 545553    epsilon: 0.1178030800
0848681      steps: 421    lr: 2.5600000000000001e-06    evaluation reward:
7.81
episode: 2084    score: 6.0    memory length: 545931    epsilon: 0.1170546400
0848732      steps: 378    lr: 2.5600000000000001e-06    evaluation reward:
7.78
episode: 2085    score: 7.0    memory length: 546354    epsilon: 0.1162171000
084879      steps: 423    lr: 2.5600000000000001e-06    evaluation reward: 7
.79
episode: 2086    score: 9.0    memory length: 546792    epsilon: 0.1153498600
```

0848849	steps: 438	lr: 2.5600000000000001e-06	evaluation reward:
7.79			
episode: 2087	score: 14.0	memory length: 547448	epsilon: 0.114050980
00848937	steps: 656	lr: 2.5600000000000001e-06	evaluation reward:
7.84			
episode: 2088	score: 9.0	memory length: 547943	epsilon: 0.1130708800
0849004	steps: 495	lr: 2.5600000000000001e-06	evaluation reward:
7.84			
episode: 2089	score: 7.0	memory length: 548349	epsilon: 0.1122670000
0849059	steps: 406	lr: 2.5600000000000001e-06	evaluation reward:
7.83			
episode: 2090	score: 16.0	memory length: 548963	epsilon: 0.111051280
00849142	steps: 614	lr: 2.5600000000000001e-06	evaluation reward:
7.92			
episode: 2091	score: 5.0	memory length: 549269	epsilon: 0.1104454000
0849183	steps: 306	lr: 2.5600000000000001e-06	evaluation reward:
7.87			
episode: 2092	score: 9.0	memory length: 549757	epsilon: 0.1094791600
0849249	steps: 488	lr: 2.5600000000000001e-06	evaluation reward:
7.88			
episode: 2093	score: 8.0	memory length: 550202	epsilon: 0.1085980600
0849309	steps: 445	lr: 2.5600000000000001e-06	evaluation reward:
7.86			
episode: 2094	score: 6.0	memory length: 550556	epsilon: 0.1078971400
0849357	steps: 354	lr: 2.5600000000000001e-06	evaluation reward:
7.83			
episode: 2095	score: 6.0	memory length: 550912	epsilon: 0.1071922600
0849405	steps: 356	lr: 2.5600000000000001e-06	evaluation reward:
7.84			
episode: 2096	score: 5.0	memory length: 551198	epsilon: 0.1066259800
0849444	steps: 286	lr: 2.5600000000000001e-06	evaluation reward:
7.85			
episode: 2097	score: 7.0	memory length: 551607	epsilon: 0.1058161600
0849499	steps: 409	lr: 2.5600000000000001e-06	evaluation reward:
7.83			
episode: 2098	score: 6.0	memory length: 551923	epsilon: 0.1051904800
0849542	steps: 316	lr: 2.5600000000000001e-06	evaluation reward:
7.79			
episode: 2099	score: 8.0	memory length: 552403	epsilon: 0.1042400800
0849606	steps: 480	lr: 2.5600000000000001e-06	evaluation reward:
7.77			
episode: 2100	score: 9.0	memory length: 552848	epsilon: 0.1033589800
0849666	steps: 445	lr: 2.5600000000000001e-06	evaluation reward:
7.79			
episode: 2101	score: 5.0	memory length: 553140	epsilon: 0.1027808200
0849706	steps: 292	lr: 2.5600000000000001e-06	evaluation reward:
7.75			
episode: 2102	score: 5.0	memory length: 553446	epsilon: 0.1021749400
0849747	steps: 306	lr: 2.5600000000000001e-06	evaluation reward:
7.71			
episode: 2103	score: 13.0	memory length: 554013	epsilon: 0.101052280
00849824	steps: 567	lr: 2.5600000000000001e-06	evaluation reward:
7.77			

episode: 2104	score: 8.0	memory length: 554464	epsilon: 0.1001593000
0849885	steps: 451	lr: 2.5600000000000001e-06	evaluation reward:
7.79			
episode: 2105	score: 10.0	memory length: 554945	epsilon: 0.099206920
0084995	steps: 481	lr: 2.5600000000000001e-06	evaluation reward:
7.82			
episode: 2106	score: 7.0	memory length: 555316	epsilon: 0.0984723400
085	steps: 371	lr: 2.5600000000000001e-06	evaluation reward: 7.83
episode: 2107	score: 5.0	memory length: 555622	epsilon: 0.0978664600
0850041	steps: 306	lr: 2.5600000000000001e-06	evaluation reward:
7.78			
episode: 2108	score: 5.0	memory length: 555932	epsilon: 0.0972526600
0850083	steps: 310	lr: 2.5600000000000001e-06	evaluation reward:
7.75			
episode: 2109	score: 3.0	memory length: 556145	epsilon: 0.0968309200
0850112	steps: 213	lr: 2.5600000000000001e-06	evaluation reward:
7.67			
episode: 2110	score: 14.0	memory length: 556760	epsilon: 0.095613220
00850195	steps: 615	lr: 2.5600000000000001e-06	evaluation reward:
7.7			
episode: 2111	score: 11.0	memory length: 557270	epsilon: 0.094603420
00850264	steps: 510	lr: 2.5600000000000001e-06	evaluation reward:
7.73			
episode: 2112	score: 11.0	memory length: 557817	epsilon: 0.093520360
00850338	steps: 547	lr: 2.5600000000000001e-06	evaluation reward:
7.78			
episode: 2113	score: 12.0	memory length: 558301	epsilon: 0.092562040
00850403	steps: 484	lr: 2.5600000000000001e-06	evaluation reward:
7.81			
episode: 2114	score: 8.0	memory length: 558757	epsilon: 0.0916591600
0850464	steps: 456	lr: 2.5600000000000001e-06	evaluation reward:
7.82			
episode: 2115	score: 8.0	memory length: 559175	epsilon: 0.0908315200
0850521	steps: 418	lr: 2.5600000000000001e-06	evaluation reward:
7.81			
episode: 2116	score: 10.0	memory length: 559717	epsilon: 0.089758360
00850594	steps: 542	lr: 2.5600000000000001e-06	evaluation reward:
7.86			
episode: 2117	score: 4.0	memory length: 559993	epsilon: 0.0892118800
0850631	steps: 276	lr: 2.5600000000000001e-06	evaluation reward:
7.84			
episode: 2118	score: 7.0	memory length: 560361	epsilon: 0.0884832400
0850681	steps: 368	lr: 2.5600000000000001e-06	evaluation reward:
7.87			
episode: 2119	score: 6.0	memory length: 560677	epsilon: 0.0878575600
0850724	steps: 316	lr: 2.5600000000000001e-06	evaluation reward:
7.85			
episode: 2120	score: 8.0	memory length: 561107	epsilon: 0.0870061600
0850782	steps: 430	lr: 2.5600000000000001e-06	evaluation reward:
7.86			
episode: 2121	score: 7.0	memory length: 561493	epsilon: 0.0862418800
0850834	steps: 386	lr: 2.5600000000000001e-06	evaluation reward:
7.83			

episode: 2122	score: 4.0	memory length: 561770	epsilon: 0.0856934200
0850871	steps: 277	lr: 2.560000000000001e-06	evaluation reward:
7.78			
episode: 2123	score: 6.0	memory length: 562144	epsilon: 0.0849529000
0850922	steps: 374	lr: 2.560000000000001e-06	evaluation reward:
7.79			
episode: 2124	score: 8.0	memory length: 562577	epsilon: 0.0840955600
085098	steps: 433	lr: 2.560000000000001e-06	evaluation reward: 7
.82			
episode: 2125	score: 3.0	memory length: 562826	epsilon: 0.0836025400
0851014	steps: 249	lr: 2.560000000000001e-06	evaluation reward:
7.78			
episode: 2126	score: 5.0	memory length: 563131	epsilon: 0.0829986400
0851055	steps: 305	lr: 2.560000000000001e-06	evaluation reward:
7.76			
episode: 2127	score: 11.0	memory length: 563703	epsilon: 0.081866080
00851132	steps: 572	lr: 2.560000000000001e-06	evaluation reward:
7.79			
episode: 2128	score: 11.0	memory length: 564220	epsilon: 0.080842420
00851202	steps: 517	lr: 2.560000000000001e-06	evaluation reward:
7.82			
episode: 2129	score: 3.0	memory length: 564447	epsilon: 0.0803929600
0851233	steps: 227	lr: 2.560000000000001e-06	evaluation reward:
7.74			
episode: 2130	score: 8.0	memory length: 564832	epsilon: 0.0796306600
0851285	steps: 385	lr: 2.560000000000001e-06	evaluation reward:
7.76			
episode: 2131	score: 8.0	memory length: 565276	epsilon: 0.0787515400
0851345	steps: 444	lr: 2.560000000000001e-06	evaluation reward:
7.77			
episode: 2132	score: 7.0	memory length: 565671	epsilon: 0.0779694400
0851398	steps: 395	lr: 2.560000000000001e-06	evaluation reward:
7.77			
episode: 2133	score: 8.0	memory length: 566121	epsilon: 0.0770784400
0851459	steps: 450	lr: 2.560000000000001e-06	evaluation reward:
7.81			
episode: 2134	score: 11.0	memory length: 566669	epsilon: 0.075993400
00851533	steps: 548	lr: 2.560000000000001e-06	evaluation reward:
7.85			
episode: 2135	score: 8.0	memory length: 567091	epsilon: 0.0751578400
085159	steps: 422	lr: 2.560000000000001e-06	evaluation reward: 7
.84			
episode: 2136	score: 4.0	memory length: 567352	epsilon: 0.0746410600
0851625	steps: 261	lr: 2.560000000000001e-06	evaluation reward:
7.81			
episode: 2137	score: 8.0	memory length: 567788	epsilon: 0.0737777800
0851684	steps: 436	lr: 2.560000000000001e-06	evaluation reward:
7.82			
episode: 2138	score: 14.0	memory length: 568351	epsilon: 0.072663040
0085176	steps: 563	lr: 2.560000000000001e-06	evaluation reward:
7.87			
episode: 2139	score: 11.0	memory length: 568903	epsilon: 0.071570080
00851835	steps: 552	lr: 2.560000000000001e-06	evaluation reward:

7.91
episode: 2140 score: 7.0 memory length: 569289 epsilon: 0.0708058000
0851887 steps: 386 lr: 2.560000000000001e-06 evaluation reward:
7.88
episode: 2141 score: 5.0 memory length: 569579 epsilon: 0.0702316000
0851926 steps: 290 lr: 2.560000000000001e-06 evaluation reward:
7.84
episode: 2142 score: 11.0 memory length: 570092 epsilon: 0.069215860
00851995 steps: 513 lr: 2.560000000000001e-06 evaluation reward:
7.87
episode: 2143 score: 10.0 memory length: 570522 epsilon: 0.068364460
00852053 steps: 430 lr: 2.560000000000001e-06 evaluation reward:
7.93
episode: 2144 score: 10.0 memory length: 570996 epsilon: 0.067425940
00852117 steps: 474 lr: 2.560000000000001e-06 evaluation reward:
7.95
episode: 2145 score: 14.0 memory length: 571543 epsilon: 0.066342880
00852191 steps: 547 lr: 2.560000000000001e-06 evaluation reward:
8.06
episode: 2146 score: 12.0 memory length: 571981 epsilon: 0.065475640
0085225 steps: 438 lr: 2.560000000000001e-06 evaluation reward:
8.07
episode: 2147 score: 9.0 memory length: 572455 epsilon: 0.0645371200
0852314 steps: 474 lr: 2.560000000000001e-06 evaluation reward:
8.13
episode: 2148 score: 3.0 memory length: 572705 epsilon: 0.0640421200
0852348 steps: 250 lr: 2.560000000000001e-06 evaluation reward:
8.0
episode: 2149 score: 6.0 memory length: 573060 epsilon: 0.0633392200
0852396 steps: 355 lr: 2.560000000000001e-06 evaluation reward:
8.0
episode: 2150 score: 10.0 memory length: 573537 epsilon: 0.062394760
008524605 steps: 477 lr: 2.560000000000001e-06 evaluation reward:
: 8.0
episode: 2151 score: 9.0 memory length: 574025 epsilon: 0.0614285200
08525264 steps: 488 lr: 2.560000000000001e-06 evaluation reward:
8.0
episode: 2152 score: 11.0 memory length: 574535 epsilon: 0.060418720
00852595 steps: 510 lr: 2.560000000000001e-06 evaluation reward:
8.06
episode: 2153 score: 8.0 memory length: 574965 epsilon: 0.0595673200
08526534 steps: 430 lr: 2.560000000000001e-06 evaluation reward:
8.02
episode: 2154 score: 5.0 memory length: 575256 epsilon: 0.0589911400
0852693 steps: 291 lr: 2.560000000000001e-06 evaluation reward:
8.03
episode: 2155 score: 9.0 memory length: 575755 epsilon: 0.0580031200
085276 steps: 499 lr: 2.560000000000001e-06 evaluation reward: 8
.01
episode: 2156 score: 6.0 memory length: 576107 epsilon: 0.0573061600
08528076 steps: 352 lr: 2.560000000000001e-06 evaluation reward:
7.95
episode: 2157 score: 5.0 memory length: 576453 epsilon: 0.0566210800

0852854	steps: 346	lr: 2.5600000000000001e-06	evaluation reward:
7.89			
episode: 2158	score: 9.0	memory length: 576877	epsilon: 0.0557815600
08529116	steps: 424	lr: 2.5600000000000001e-06	evaluation reward:
7.94			
episode: 2159	score: 7.0	memory length: 577268	epsilon: 0.0550073800
08529644	steps: 391	lr: 2.5600000000000001e-06	evaluation reward:
7.97			
episode: 2160	score: 9.0	memory length: 577741	epsilon: 0.0540708400
0853028	steps: 473	lr: 2.5600000000000001e-06	evaluation reward:
7.99			
episode: 2161	score: 9.0	memory length: 578188	epsilon: 0.0531857800
08530886	steps: 447	lr: 2.5600000000000001e-06	evaluation reward:
8.01			
episode: 2162	score: 9.0	memory length: 578640	epsilon: 0.0522908200
085315	steps: 452	lr: 2.5600000000000001e-06	evaluation reward: 8
.04			
episode: 2163	score: 9.0	memory length: 579132	epsilon: 0.0513166600
0853216	steps: 492	lr: 2.5600000000000001e-06	evaluation reward:
8.04			
episode: 2164	score: 8.0	memory length: 579576	epsilon: 0.0504375400
0853276	steps: 444	lr: 2.5600000000000001e-06	evaluation reward:
8.06			
episode: 2165	score: 9.0	memory length: 580070	epsilon: 0.0494594200
0853343	steps: 494	lr: 2.5600000000000001e-06	evaluation reward:
8.06			
episode: 2166	score: 15.0	memory length: 580586	epsilon: 0.048437740
008534125	steps: 516	lr: 2.5600000000000001e-06	evaluation reward
: 8.12			
episode: 2167	score: 11.0	memory length: 581081	epsilon: 0.047457640
00853479	steps: 495	lr: 2.5600000000000001e-06	evaluation reward:
8.11			
episode: 2168	score: 12.0	memory length: 581521	epsilon: 0.046586440
00853539	steps: 440	lr: 2.5600000000000001e-06	evaluation reward:
8.14			
episode: 2169	score: 9.0	memory length: 581995	epsilon: 0.0456479200
0853603	steps: 474	lr: 2.5600000000000001e-06	evaluation reward:
8.18			
episode: 2170	score: 10.0	memory length: 582517	epsilon: 0.044614360
00853673	steps: 522	lr: 2.5600000000000001e-06	evaluation reward:
8.2			
episode: 2171	score: 9.0	memory length: 582969	epsilon: 0.0437194000
0853734	steps: 452	lr: 2.5600000000000001e-06	evaluation reward:
8.2			
episode: 2172	score: 6.0	memory length: 583326	epsilon: 0.0430125400
08537825	steps: 357	lr: 2.5600000000000001e-06	evaluation reward:
8.19			
episode: 2173	score: 14.0	memory length: 583956	epsilon: 0.041765140
008538676	steps: 630	lr: 2.5600000000000001e-06	evaluation reward
: 8.27			
episode: 2174	score: 8.0	memory length: 584373	epsilon: 0.0409394800
0853924	steps: 417	lr: 2.5600000000000001e-06	evaluation reward:
8.27			

episode: 2175	score: 15.0	memory length: 584964	epsilon: 0.039769300
00854004	steps: 591	lr: 2.560000000000001e-06	evaluation reward:
8.34			
episode: 2176	score: 6.0	memory length: 585324	epsilon: 0.0390565000
0854052	steps: 360	lr: 2.560000000000001e-06	evaluation reward:
8.29			
episode: 2177	score: 11.0	memory length: 585915	epsilon: 0.037886320
00854132	steps: 591	lr: 2.560000000000001e-06	evaluation reward:
8.33			
episode: 2178	score: 9.0	memory length: 586406	epsilon: 0.0369141400
08541985	steps: 491	lr: 2.560000000000001e-06	evaluation reward:
8.34			
episode: 2179	score: 6.0	memory length: 586785	epsilon: 0.0361637200
08542496	steps: 379	lr: 2.560000000000001e-06	evaluation reward:
8.33			
episode: 2180	score: 7.0	memory length: 587195	epsilon: 0.0353519200
0854305	steps: 410	lr: 2.560000000000001e-06	evaluation reward:
8.3			
episode: 2181	score: 13.0	memory length: 587767	epsilon: 0.034219360
00854382	steps: 572	lr: 2.560000000000001e-06	evaluation reward:
8.36			
episode: 2182	score: 9.0	memory length: 588215	epsilon: 0.0333323200
0854443	steps: 448	lr: 2.560000000000001e-06	evaluation reward:
8.39			
episode: 2183	score: 9.0	memory length: 588708	epsilon: 0.0323561800
0854509	steps: 493	lr: 2.560000000000001e-06	evaluation reward:
8.4			
episode: 2184	score: 7.0	memory length: 589077	epsilon: 0.0316255600
0854559	steps: 369	lr: 2.560000000000001e-06	evaluation reward:
8.41			
episode: 2185	score: 4.0	memory length: 589358	epsilon: 0.0310691800
0854597	steps: 281	lr: 2.560000000000001e-06	evaluation reward:
8.38			
episode: 2186	score: 7.0	memory length: 589726	epsilon: 0.0303405400
08546468	steps: 368	lr: 2.560000000000001e-06	evaluation reward:
8.36			
episode: 2187	score: 14.0	memory length: 590310	epsilon: 0.029184220
008547257	steps: 584	lr: 2.560000000000001e-06	evaluation reward
: 8.36			
episode: 2188	score: 7.0	memory length: 590693	epsilon: 0.0284258800
08547774	steps: 383	lr: 2.560000000000001e-06	evaluation reward:
8.34			
episode: 2189	score: 8.0	memory length: 591096	epsilon: 0.0276279400
0854832	steps: 403	lr: 2.560000000000001e-06	evaluation reward:
8.35			
episode: 2190	score: 10.0	memory length: 591611	epsilon: 0.026608240
008549014	steps: 515	lr: 2.560000000000001e-06	evaluation reward
: 8.29			
episode: 2191	score: 6.0	memory length: 591935	epsilon: 0.0259667200
0854945	steps: 324	lr: 2.560000000000001e-06	evaluation reward:
8.3			
episode: 2192	score: 12.0	memory length: 592468	epsilon: 0.024911380
00855017	steps: 533	lr: 2.560000000000001e-06	evaluation reward:

8.33
episode: 2193 score: 8.0 memory length: 592854 epsilon: 0.0241471000
08550693 steps: 386 lr: 2.560000000000001e-06 evaluation reward:
8.33
episode: 2194 score: 5.0 memory length: 593161 epsilon: 0.0235392400
08551107 steps: 307 lr: 2.560000000000001e-06 evaluation reward:
8.32
episode: 2195 score: 11.0 memory length: 593720 epsilon: 0.022432420
008551862 steps: 559 lr: 2.560000000000001e-06 evaluation reward
: 8.37
episode: 2196 score: 5.0 memory length: 594042 epsilon: 0.0217948600
08552297 steps: 322 lr: 2.560000000000001e-06 evaluation reward:
8.37
episode: 2197 score: 8.0 memory length: 594498 epsilon: 0.0208919800
08552913 steps: 456 lr: 2.560000000000001e-06 evaluation reward:
8.38
episode: 2198 score: 11.0 memory length: 595009 epsilon: 0.019880200
008553603 steps: 511 lr: 2.560000000000001e-06 evaluation reward
: 8.43
episode: 2199 score: 3.0 memory length: 595259 epsilon: 0.0193852000
0855394 steps: 250 lr: 2.560000000000001e-06 evaluation reward:
8.38
episode: 2200 score: 5.0 memory length: 595567 epsilon: 0.0187753600
08554356 steps: 308 lr: 2.560000000000001e-06 evaluation reward:
8.34
episode: 2201 score: 6.0 memory length: 595924 epsilon: 0.0180685000
0855484 steps: 357 lr: 2.560000000000001e-06 evaluation reward:
8.35
episode: 2202 score: 9.0 memory length: 596370 epsilon: 0.0171854200
0855544 steps: 446 lr: 2.560000000000001e-06 evaluation reward:
8.39
episode: 2203 score: 10.0 memory length: 596859 epsilon: 0.016217200
0085561 steps: 489 lr: 2.560000000000001e-06 evaluation reward:
8.36
episode: 2204 score: 9.0 memory length: 597296 epsilon: 0.0153519400
08556452 steps: 437 lr: 2.560000000000001e-06 evaluation reward:
8.37
episode: 2205 score: 10.0 memory length: 597785 epsilon: 0.014383720
008556264 steps: 489 lr: 2.560000000000001e-06 evaluation reward
: 8.37
episode: 2206 score: 8.0 memory length: 598208 epsilon: 0.0135461800
08556102 steps: 423 lr: 2.560000000000001e-06 evaluation reward:
8.38
episode: 2207 score: 11.0 memory length: 598764 epsilon: 0.012445300
008555888 steps: 556 lr: 2.560000000000001e-06 evaluation reward
: 8.44
episode: 2208 score: 8.0 memory length: 599206 epsilon: 0.0115701400
08555718 steps: 442 lr: 2.560000000000001e-06 evaluation reward:
8.47
episode: 2209 score: 10.0 memory length: 599737 epsilon: 0.010518760
008555514 steps: 531 lr: 2.560000000000001e-06 evaluation reward
: 8.54
episode: 2210 score: 8.0 memory length: 600158 epsilon: 0.0099980200

08555413 steps: 421 lr: 1.0240000000000005e-06 evaluation reward : 8.48
episode: 2211 score: 12.0 memory length: 600753 epsilon: 0.009998020
08555413 steps: 595 lr: 1.0240000000000005e-06 evaluation reward : 8.49
episode: 2212 score: 9.0 memory length: 601234 epsilon: 0.0099980200
08555413 steps: 481 lr: 1.0240000000000005e-06 evaluation reward : 8.47
episode: 2213 score: 8.0 memory length: 601706 epsilon: 0.0099980200
08555413 steps: 472 lr: 1.0240000000000005e-06 evaluation reward : 8.43
episode: 2214 score: 8.0 memory length: 602116 epsilon: 0.0099980200
08555413 steps: 410 lr: 1.0240000000000005e-06 evaluation reward : 8.43
episode: 2215 score: 7.0 memory length: 602522 epsilon: 0.0099980200
08555413 steps: 406 lr: 1.0240000000000005e-06 evaluation reward : 8.42
episode: 2216 score: 7.0 memory length: 602950 epsilon: 0.0099980200
08555413 steps: 428 lr: 1.0240000000000005e-06 evaluation reward : 8.39
episode: 2217 score: 11.0 memory length: 603488 epsilon: 0.009998020
08555413 steps: 538 lr: 1.0240000000000005e-06 evaluation reward : 8.46
episode: 2218 score: 6.0 memory length: 603863 epsilon: 0.0099980200
08555413 steps: 375 lr: 1.0240000000000005e-06 evaluation reward : 8.45
episode: 2219 score: 6.0 memory length: 604238 epsilon: 0.0099980200
08555413 steps: 375 lr: 1.0240000000000005e-06 evaluation reward : 8.45
episode: 2220 score: 8.0 memory length: 604673 epsilon: 0.0099980200
08555413 steps: 435 lr: 1.0240000000000005e-06 evaluation reward : 8.45
episode: 2221 score: 9.0 memory length: 605166 epsilon: 0.0099980200
08555413 steps: 493 lr: 1.0240000000000005e-06 evaluation reward : 8.47
episode: 2222 score: 6.0 memory length: 605503 epsilon: 0.0099980200
08555413 steps: 337 lr: 1.0240000000000005e-06 evaluation reward : 8.49
episode: 2223 score: 9.0 memory length: 606005 epsilon: 0.0099980200
08555413 steps: 502 lr: 1.0240000000000005e-06 evaluation reward : 8.52
episode: 2224 score: 5.0 memory length: 606335 epsilon: 0.0099980200
08555413 steps: 330 lr: 1.0240000000000005e-06 evaluation reward : 8.49
episode: 2225 score: 15.0 memory length: 606952 epsilon: 0.009998020
08555413 steps: 617 lr: 1.0240000000000005e-06 evaluation reward : 8.61
episode: 2226 score: 8.0 memory length: 607371 epsilon: 0.0099980200
08555413 steps: 419 lr: 1.0240000000000005e-06 evaluation reward : 8.64
episode: 2227 score: 10.0 memory length: 607806 epsilon: 0.009998020
08555413 steps: 435 lr: 1.0240000000000005e-06 evaluation reward : 8.63

episode: 2228 score: 8.0 memory length: 608161 epsilon: 0.0099980200
08555413 steps: 355 lr: 1.0240000000000005e-06 evaluation reward
: 8.6

episode: 2229 score: 7.0 memory length: 608544 epsilon: 0.0099980200
08555413 steps: 383 lr: 1.0240000000000005e-06 evaluation reward
: 8.64

episode: 2230 score: 6.0 memory length: 608906 epsilon: 0.0099980200
08555413 steps: 362 lr: 1.0240000000000005e-06 evaluation reward
: 8.62

episode: 2231 score: 13.0 memory length: 609567 epsilon: 0.0099980200
08555413 steps: 661 lr: 1.0240000000000005e-06 evaluation reward
: 8.67

episode: 2232 score: 10.0 memory length: 610100 epsilon: 0.0099980200
08555413 steps: 533 lr: 1.0240000000000005e-06 evaluation reward
: 8.7

episode: 2233 score: 14.0 memory length: 610726 epsilon: 0.0099980200
08555413 steps: 626 lr: 1.0240000000000005e-06 evaluation reward
: 8.76

episode: 2234 score: 9.0 memory length: 611218 epsilon: 0.0099980200
08555413 steps: 492 lr: 1.0240000000000005e-06 evaluation reward
: 8.74

episode: 2235 score: 7.0 memory length: 611603 epsilon: 0.0099980200
08555413 steps: 385 lr: 1.0240000000000005e-06 evaluation reward
: 8.73

episode: 2236 score: 9.0 memory length: 612027 epsilon: 0.0099980200
08555413 steps: 424 lr: 1.0240000000000005e-06 evaluation reward
: 8.78

episode: 2237 score: 6.0 memory length: 612381 epsilon: 0.0099980200
08555413 steps: 354 lr: 1.0240000000000005e-06 evaluation reward
: 8.76

episode: 2238 score: 8.0 memory length: 612758 epsilon: 0.0099980200
08555413 steps: 377 lr: 1.0240000000000005e-06 evaluation reward
: 8.7

episode: 2239 score: 8.0 memory length: 613135 epsilon: 0.0099980200
08555413 steps: 377 lr: 1.0240000000000005e-06 evaluation reward
: 8.67

episode: 2240 score: 11.0 memory length: 613669 epsilon: 0.0099980200
08555413 steps: 534 lr: 1.0240000000000005e-06 evaluation reward
: 8.71

episode: 2241 score: 7.0 memory length: 614018 epsilon: 0.0099980200
08555413 steps: 349 lr: 1.0240000000000005e-06 evaluation reward
: 8.73

episode: 2242 score: 9.0 memory length: 614484 epsilon: 0.0099980200
08555413 steps: 466 lr: 1.0240000000000005e-06 evaluation reward
: 8.71

episode: 2243 score: 9.0 memory length: 614967 epsilon: 0.0099980200
08555413 steps: 483 lr: 1.0240000000000005e-06 evaluation reward
: 8.7

episode: 2244 score: 9.0 memory length: 615415 epsilon: 0.0099980200
08555413 steps: 448 lr: 1.0240000000000005e-06 evaluation reward
: 8.69

episode: 2245 score: 10.0 memory length: 615937 epsilon: 0.0099980200
08555413 steps: 522 lr: 1.0240000000000005e-06 evaluation reward

d: 8.65
episode: 2246 score: 10.0 memory length: 616473 epsilon: 0.009998020
008555413 steps: 536 lr: 1.0240000000000005e-06 evaluation reward
d: 8.63
episode: 2247 score: 13.0 memory length: 617084 epsilon: 0.009998020
008555413 steps: 611 lr: 1.0240000000000005e-06 evaluation reward
d: 8.67
episode: 2248 score: 10.0 memory length: 617560 epsilon: 0.009998020
008555413 steps: 476 lr: 1.0240000000000005e-06 evaluation reward
d: 8.74
episode: 2249 score: 13.0 memory length: 618159 epsilon: 0.009998020
008555413 steps: 599 lr: 1.0240000000000005e-06 evaluation reward
d: 8.81
episode: 2250 score: 10.0 memory length: 618689 epsilon: 0.009998020
008555413 steps: 530 lr: 1.0240000000000005e-06 evaluation reward
d: 8.81
episode: 2251 score: 11.0 memory length: 619221 epsilon: 0.009998020
008555413 steps: 532 lr: 1.0240000000000005e-06 evaluation reward
d: 8.83
episode: 2252 score: 8.0 memory length: 619628 epsilon: 0.0099980200
08555413 steps: 407 lr: 1.0240000000000005e-06 evaluation reward
: 8.8
episode: 2253 score: 6.0 memory length: 619939 epsilon: 0.0099980200
08555413 steps: 311 lr: 1.0240000000000005e-06 evaluation reward
: 8.78
episode: 2254 score: 13.0 memory length: 620420 epsilon: 0.009998020
008555413 steps: 481 lr: 1.0240000000000005e-06 evaluation reward
d: 8.86
episode: 2255 score: 11.0 memory length: 620966 epsilon: 0.009998020
008555413 steps: 546 lr: 1.0240000000000005e-06 evaluation reward
d: 8.88
episode: 2256 score: 6.0 memory length: 621305 epsilon: 0.0099980200
08555413 steps: 339 lr: 1.0240000000000005e-06 evaluation reward
: 8.88
episode: 2257 score: 11.0 memory length: 621833 epsilon: 0.009998020
008555413 steps: 528 lr: 1.0240000000000005e-06 evaluation reward
d: 8.94
episode: 2258 score: 4.0 memory length: 622110 epsilon: 0.0099980200
08555413 steps: 277 lr: 1.0240000000000005e-06 evaluation reward
: 8.89
episode: 2259 score: 8.0 memory length: 622551 epsilon: 0.0099980200
08555413 steps: 441 lr: 1.0240000000000005e-06 evaluation reward
: 8.9
episode: 2260 score: 8.0 memory length: 622993 epsilon: 0.0099980200
08555413 steps: 442 lr: 1.0240000000000005e-06 evaluation reward
: 8.89
episode: 2261 score: 5.0 memory length: 623319 epsilon: 0.0099980200
08555413 steps: 326 lr: 1.0240000000000005e-06 evaluation reward
: 8.85
episode: 2262 score: 6.0 memory length: 623683 epsilon: 0.0099980200
08555413 steps: 364 lr: 1.0240000000000005e-06 evaluation reward
: 8.82
episode: 2263 score: 6.0 memory length: 624034 epsilon: 0.0099980200

08555413 steps: 351 lr: 1.0240000000000005e-06 evaluation reward : 8.79
episode: 2264 score: 11.0 memory length: 624542 epsilon: 0.009998020
08555413 steps: 508 lr: 1.0240000000000005e-06 evaluation reward : 8.82
episode: 2265 score: 8.0 memory length: 625006 epsilon: 0.0099980200
08555413 steps: 464 lr: 1.0240000000000005e-06 evaluation reward : 8.81
episode: 2266 score: 13.0 memory length: 625466 epsilon: 0.009998020
08555413 steps: 460 lr: 1.0240000000000005e-06 evaluation reward : 8.79
episode: 2267 score: 7.0 memory length: 625872 epsilon: 0.0099980200
08555413 steps: 406 lr: 1.0240000000000005e-06 evaluation reward : 8.75
episode: 2268 score: 9.0 memory length: 626321 epsilon: 0.0099980200
08555413 steps: 449 lr: 1.0240000000000005e-06 evaluation reward : 8.72
episode: 2269 score: 15.0 memory length: 626742 epsilon: 0.009998020
08555413 steps: 421 lr: 1.0240000000000005e-06 evaluation reward : 8.78
episode: 2270 score: 9.0 memory length: 627210 epsilon: 0.0099980200
08555413 steps: 468 lr: 1.0240000000000005e-06 evaluation reward : 8.77
episode: 2271 score: 12.0 memory length: 627774 epsilon: 0.009998020
08555413 steps: 564 lr: 1.0240000000000005e-06 evaluation reward : 8.8
episode: 2272 score: 11.0 memory length: 628199 epsilon: 0.009998020
08555413 steps: 425 lr: 1.0240000000000005e-06 evaluation reward : 8.85
episode: 2273 score: 10.0 memory length: 628703 epsilon: 0.009998020
08555413 steps: 504 lr: 1.0240000000000005e-06 evaluation reward : 8.81
episode: 2274 score: 5.0 memory length: 629032 epsilon: 0.0099980200
08555413 steps: 329 lr: 1.0240000000000005e-06 evaluation reward : 8.78
episode: 2275 score: 8.0 memory length: 629509 epsilon: 0.0099980200
08555413 steps: 477 lr: 1.0240000000000005e-06 evaluation reward : 8.71
episode: 2276 score: 6.0 memory length: 629883 epsilon: 0.0099980200
08555413 steps: 374 lr: 1.0240000000000005e-06 evaluation reward : 8.71
episode: 2277 score: 8.0 memory length: 630348 epsilon: 0.0099980200
08555413 steps: 465 lr: 1.0240000000000005e-06 evaluation reward : 8.68
episode: 2278 score: 10.0 memory length: 630827 epsilon: 0.009998020
08555413 steps: 479 lr: 1.0240000000000005e-06 evaluation reward : 8.69
episode: 2279 score: 10.0 memory length: 631310 epsilon: 0.009998020
08555413 steps: 483 lr: 1.0240000000000005e-06 evaluation reward : 8.73
episode: 2280 score: 8.0 memory length: 631761 epsilon: 0.0099980200
08555413 steps: 451 lr: 1.0240000000000005e-06 evaluation reward : 8.74

episode: 2281 score: 13.0 memory length: 632393 epsilon: 0.009998020
008555413 steps: 632 lr: 1.0240000000000005e-06 evaluation reward:
8.74

episode: 2282 score: 8.0 memory length: 632804 epsilon: 0.0099980200
008555413 steps: 411 lr: 1.0240000000000005e-06 evaluation reward
: 8.73

episode: 2283 score: 8.0 memory length: 633218 epsilon: 0.0099980200
008555413 steps: 414 lr: 1.0240000000000005e-06 evaluation reward
: 8.72

episode: 2284 score: 10.0 memory length: 633718 epsilon: 0.009998020
008555413 steps: 500 lr: 1.0240000000000005e-06 evaluation reward:
8.75

episode: 2285 score: 7.0 memory length: 634124 epsilon: 0.0099980200
008555413 steps: 406 lr: 1.0240000000000005e-06 evaluation reward
: 8.78

episode: 2286 score: 11.0 memory length: 634681 epsilon: 0.009998020
008555413 steps: 557 lr: 1.0240000000000005e-06 evaluation reward:
8.82

episode: 2287 score: 9.0 memory length: 635175 epsilon: 0.0099980200
008555413 steps: 494 lr: 1.0240000000000005e-06 evaluation reward
: 8.77

episode: 2288 score: 7.0 memory length: 635568 epsilon: 0.0099980200
008555413 steps: 393 lr: 1.0240000000000005e-06 evaluation reward
: 8.77

episode: 2289 score: 6.0 memory length: 635948 epsilon: 0.0099980200
008555413 steps: 380 lr: 1.0240000000000005e-06 evaluation reward
: 8.75

episode: 2290 score: 7.0 memory length: 636346 epsilon: 0.0099980200
008555413 steps: 398 lr: 1.0240000000000005e-06 evaluation reward
: 8.72

episode: 2291 score: 6.0 memory length: 636675 epsilon: 0.0099980200
008555413 steps: 329 lr: 1.0240000000000005e-06 evaluation reward
: 8.72

episode: 2292 score: 7.0 memory length: 637059 epsilon: 0.0099980200
008555413 steps: 384 lr: 1.0240000000000005e-06 evaluation reward
: 8.67

episode: 2293 score: 10.0 memory length: 637603 epsilon: 0.009998020
008555413 steps: 544 lr: 1.0240000000000005e-06 evaluation reward:
8.69

episode: 2294 score: 7.0 memory length: 637972 epsilon: 0.0099980200
008555413 steps: 369 lr: 1.0240000000000005e-06 evaluation reward
: 8.71

episode: 2295 score: 9.0 memory length: 638441 epsilon: 0.0099980200
008555413 steps: 469 lr: 1.0240000000000005e-06 evaluation reward
: 8.69

episode: 2296 score: 16.0 memory length: 639108 epsilon: 0.009998020
008555413 steps: 667 lr: 1.0240000000000005e-06 evaluation reward:
8.8

episode: 2297 score: 16.0 memory length: 639718 epsilon: 0.009998020
008555413 steps: 610 lr: 1.0240000000000005e-06 evaluation reward:
8.88

episode: 2298 score: 9.0 memory length: 640196 epsilon: 0.0099980200
008555413 steps: 478 lr: 1.0240000000000005e-06 evaluation reward

: 8.86
episode: 2299 score: 6.0 memory length: 640552 epsilon: 0.0099980200
08555413 steps: 356 lr: 1.0240000000000005e-06 evaluation reward
: 8.89
episode: 2300 score: 11.0 memory length: 641122 epsilon: 0.0099980200
08555413 steps: 570 lr: 1.0240000000000005e-06 evaluation reward
d: 8.95
episode: 2301 score: 15.0 memory length: 641706 epsilon: 0.0099980200
08555413 steps: 584 lr: 1.0240000000000005e-06 evaluation reward
d: 9.04
episode: 2302 score: 7.0 memory length: 642111 epsilon: 0.0099980200
08555413 steps: 405 lr: 1.0240000000000005e-06 evaluation reward
: 9.02
episode: 2303 score: 5.0 memory length: 642422 epsilon: 0.0099980200
08555413 steps: 311 lr: 1.0240000000000005e-06 evaluation reward
: 8.97
episode: 2304 score: 12.0 memory length: 642918 epsilon: 0.0099980200
08555413 steps: 496 lr: 1.0240000000000005e-06 evaluation reward
d: 9.0
episode: 2305 score: 7.0 memory length: 643301 epsilon: 0.0099980200
08555413 steps: 383 lr: 1.0240000000000005e-06 evaluation reward
: 8.97
episode: 2306 score: 11.0 memory length: 643869 epsilon: 0.0099980200
08555413 steps: 568 lr: 1.0240000000000005e-06 evaluation reward
d: 9.0
episode: 2307 score: 12.0 memory length: 644433 epsilon: 0.0099980200
08555413 steps: 564 lr: 1.0240000000000005e-06 evaluation reward
d: 9.01
episode: 2308 score: 7.0 memory length: 644841 epsilon: 0.0099980200
08555413 steps: 408 lr: 1.0240000000000005e-06 evaluation reward
: 9.0
episode: 2309 score: 11.0 memory length: 645244 epsilon: 0.0099980200
08555413 steps: 403 lr: 1.0240000000000005e-06 evaluation reward
d: 9.01
episode: 2310 score: 6.0 memory length: 645601 epsilon: 0.0099980200
08555413 steps: 357 lr: 1.0240000000000005e-06 evaluation reward
: 8.99
episode: 2311 score: 10.0 memory length: 646055 epsilon: 0.0099980200
08555413 steps: 454 lr: 1.0240000000000005e-06 evaluation reward
d: 8.97
episode: 2312 score: 7.0 memory length: 646460 epsilon: 0.0099980200
08555413 steps: 405 lr: 1.0240000000000005e-06 evaluation reward
: 8.95
episode: 2313 score: 10.0 memory length: 646940 epsilon: 0.0099980200
08555413 steps: 480 lr: 1.0240000000000005e-06 evaluation reward
d: 8.97
episode: 2314 score: 8.0 memory length: 647373 epsilon: 0.0099980200
08555413 steps: 433 lr: 1.0240000000000005e-06 evaluation reward
: 8.97
episode: 2315 score: 7.0 memory length: 647779 epsilon: 0.0099980200
08555413 steps: 406 lr: 1.0240000000000005e-06 evaluation reward
: 8.97
episode: 2316 score: 7.0 memory length: 648189 epsilon: 0.0099980200

008555413 steps: 410 lr: 1.0240000000000005e-06 evaluation reward : 8.97
episode: 2317 score: 17.0 memory length: 648840 epsilon: 0.009998020
008555413 steps: 651 lr: 1.0240000000000005e-06 evaluation reward : 9.03
episode: 2318 score: 12.0 memory length: 649446 epsilon: 0.009998020
008555413 steps: 606 lr: 1.0240000000000005e-06 evaluation reward : 9.09
episode: 2319 score: 6.0 memory length: 649762 epsilon: 0.009998020
008555413 steps: 316 lr: 1.0240000000000005e-06 evaluation reward : 9.09
episode: 2320 score: 14.0 memory length: 650377 epsilon: 0.009998020
008555413 steps: 615 lr: 1.0240000000000005e-06 evaluation reward : 9.15
episode: 2321 score: 13.0 memory length: 650972 epsilon: 0.009998020
008555413 steps: 595 lr: 1.0240000000000005e-06 evaluation reward : 9.19
episode: 2322 score: 8.0 memory length: 651398 epsilon: 0.009998020
008555413 steps: 426 lr: 1.0240000000000005e-06 evaluation reward : 9.21
episode: 2323 score: 9.0 memory length: 651880 epsilon: 0.009998020
008555413 steps: 482 lr: 1.0240000000000005e-06 evaluation reward : 9.21
episode: 2324 score: 5.0 memory length: 652189 epsilon: 0.009998020
008555413 steps: 309 lr: 1.0240000000000005e-06 evaluation reward : 9.21
episode: 2325 score: 10.0 memory length: 652724 epsilon: 0.009998020
008555413 steps: 535 lr: 1.0240000000000005e-06 evaluation reward : 9.16
episode: 2326 score: 14.0 memory length: 653352 epsilon: 0.009998020
008555413 steps: 628 lr: 1.0240000000000005e-06 evaluation reward : 9.22
episode: 2327 score: 9.0 memory length: 653854 epsilon: 0.009998020
008555413 steps: 502 lr: 1.0240000000000005e-06 evaluation reward : 9.21
episode: 2328 score: 15.0 memory length: 654411 epsilon: 0.009998020
008555413 steps: 557 lr: 1.0240000000000005e-06 evaluation reward : 9.28
episode: 2329 score: 14.0 memory length: 654942 epsilon: 0.009998020
008555413 steps: 531 lr: 1.0240000000000005e-06 evaluation reward : 9.35
episode: 2330 score: 9.0 memory length: 655435 epsilon: 0.009998020
008555413 steps: 493 lr: 1.0240000000000005e-06 evaluation reward : 9.38
episode: 2331 score: 8.0 memory length: 655855 epsilon: 0.009998020
008555413 steps: 420 lr: 1.0240000000000005e-06 evaluation reward : 9.33
episode: 2332 score: 9.0 memory length: 656303 epsilon: 0.009998020
008555413 steps: 448 lr: 1.0240000000000005e-06 evaluation reward : 9.32
episode: 2333 score: 7.0 memory length: 656709 epsilon: 0.009998020
008555413 steps: 406 lr: 1.0240000000000005e-06 evaluation reward : 9.25

episode: 2334 score: 8.0 memory length: 657153 epsilon: 0.0099980200
08555413 steps: 444 lr: 1.0240000000000005e-06 evaluation reward
: 9.24

episode: 2335 score: 10.0 memory length: 657677 epsilon: 0.0099980200
08555413 steps: 524 lr: 1.0240000000000005e-06 evaluation reward
d: 9.27

episode: 2336 score: 13.0 memory length: 658272 epsilon: 0.0099980200
08555413 steps: 595 lr: 1.0240000000000005e-06 evaluation reward
d: 9.31

episode: 2337 score: 8.0 memory length: 658711 epsilon: 0.0099980200
08555413 steps: 439 lr: 1.0240000000000005e-06 evaluation reward
: 9.33

episode: 2338 score: 11.0 memory length: 659252 epsilon: 0.0099980200
08555413 steps: 541 lr: 1.0240000000000005e-06 evaluation reward
d: 9.36

episode: 2339 score: 8.0 memory length: 659586 epsilon: 0.0099980200
08555413 steps: 334 lr: 1.0240000000000005e-06 evaluation reward
: 9.36

episode: 2340 score: 10.0 memory length: 660118 epsilon: 0.0099980200
08555413 steps: 532 lr: 1.0240000000000005e-06 evaluation reward
d: 9.35

episode: 2341 score: 8.0 memory length: 660533 epsilon: 0.0099980200
08555413 steps: 415 lr: 1.0240000000000005e-06 evaluation reward
: 9.36

episode: 2342 score: 5.0 memory length: 660826 epsilon: 0.0099980200
08555413 steps: 293 lr: 1.0240000000000005e-06 evaluation reward
: 9.32

episode: 2343 score: 6.0 memory length: 661188 epsilon: 0.0099980200
08555413 steps: 362 lr: 1.0240000000000005e-06 evaluation reward
: 9.29

episode: 2344 score: 4.0 memory length: 661467 epsilon: 0.0099980200
08555413 steps: 279 lr: 1.0240000000000005e-06 evaluation reward
: 9.24

episode: 2345 score: 6.0 memory length: 661818 epsilon: 0.0099980200
08555413 steps: 351 lr: 1.0240000000000005e-06 evaluation reward
: 9.2

episode: 2346 score: 4.0 memory length: 662097 epsilon: 0.0099980200
08555413 steps: 279 lr: 1.0240000000000005e-06 evaluation reward
: 9.14

episode: 2347 score: 10.0 memory length: 662613 epsilon: 0.0099980200
08555413 steps: 516 lr: 1.0240000000000005e-06 evaluation reward
d: 9.11

episode: 2348 score: 8.0 memory length: 663033 epsilon: 0.0099980200
08555413 steps: 420 lr: 1.0240000000000005e-06 evaluation reward
: 9.09

episode: 2349 score: 8.0 memory length: 663430 epsilon: 0.0099980200
08555413 steps: 397 lr: 1.0240000000000005e-06 evaluation reward
: 9.04

episode: 2350 score: 15.0 memory length: 664053 epsilon: 0.0099980200
08555413 steps: 623 lr: 1.0240000000000005e-06 evaluation reward
d: 9.09

episode: 2351 score: 6.0 memory length: 664425 epsilon: 0.0099980200
08555413 steps: 372 lr: 1.0240000000000005e-06 evaluation reward


```
: 9.04
episode: 2352  score: 9.0  memory length: 664906  epsilon: 0.0099980200
08555413  steps: 481  lr: 1.0240000000000005e-06  evaluation reward
: 9.05
episode: 2353  score: 9.0  memory length: 665365  epsilon: 0.0099980200
08555413  steps: 459  lr: 1.0240000000000005e-06  evaluation reward
: 9.08
episode: 2354  score: 7.0  memory length: 665788  epsilon: 0.0099980200
08555413  steps: 423  lr: 1.0240000000000005e-06  evaluation reward
: 9.02
episode: 2355  score: 5.0  memory length: 666116  epsilon: 0.0099980200
08555413  steps: 328  lr: 1.0240000000000005e-06  evaluation reward
: 8.96
episode: 2356  score: 8.0  memory length: 666530  epsilon: 0.0099980200
08555413  steps: 414  lr: 1.0240000000000005e-06  evaluation reward
: 8.98
episode: 2357  score: 13.0  memory length: 667132  epsilon: 0.0099980200
08555413  steps: 602  lr: 1.0240000000000005e-06  evaluation reward
d: 9.0
episode: 2358  score: 11.0  memory length: 667655  epsilon: 0.0099980200
08555413  steps: 523  lr: 1.0240000000000005e-06  evaluation reward
d: 9.07
episode: 2359  score: 5.0  memory length: 667945  epsilon: 0.0099980200
08555413  steps: 290  lr: 1.0240000000000005e-06  evaluation reward
: 9.04
episode: 2360  score: 13.0  memory length: 668455  epsilon: 0.0099980200
08555413  steps: 510  lr: 1.0240000000000005e-06  evaluation reward
d: 9.09
episode: 2361  score: 11.0  memory length: 669005  epsilon: 0.0099980200
08555413  steps: 550  lr: 1.0240000000000005e-06  evaluation reward
d: 9.15
episode: 2362  score: 10.0  memory length: 669465  epsilon: 0.0099980200
08555413  steps: 460  lr: 1.0240000000000005e-06  evaluation reward
d: 9.19
episode: 2363  score: 11.0  memory length: 669983  epsilon: 0.0099980200
08555413  steps: 518  lr: 1.0240000000000005e-06  evaluation reward
d: 9.24
episode: 2364  score: 9.0  memory length: 670437  epsilon: 0.0099980200
08555413  steps: 454  lr: 1.0240000000000005e-06  evaluation reward
: 9.22
episode: 2365  score: 11.0  memory length: 670853  epsilon: 0.0099980200
08555413  steps: 416  lr: 1.0240000000000005e-06  evaluation reward
d: 9.25
episode: 2366  score: 7.0  memory length: 671221  epsilon: 0.0099980200
08555413  steps: 368  lr: 1.0240000000000005e-06  evaluation reward
: 9.19
episode: 2367  score: 5.0  memory length: 671492  epsilon: 0.0099980200
08555413  steps: 271  lr: 1.0240000000000005e-06  evaluation reward
: 9.17
episode: 2368  score: 9.0  memory length: 671901  epsilon: 0.0099980200
08555413  steps: 409  lr: 1.0240000000000005e-06  evaluation reward
: 9.17
episode: 2369  score: 6.0  memory length: 672223  epsilon: 0.0099980200
```

```
08555413    steps: 322    lr: 1.0240000000000005e-06    evaluation reward
: 9.08
episode: 2370    score: 9.0    memory length: 672661    epsilon: 0.0099980200
08555413    steps: 438    lr: 1.0240000000000005e-06    evaluation reward
: 9.08
episode: 2371    score: 11.0    memory length: 673183    epsilon: 0.009998020
08555413    steps: 522    lr: 1.0240000000000005e-06    evaluation reward: 9.07
episode: 2372    score: 21.0    memory length: 673950    epsilon: 0.009998020
08555413    steps: 767    lr: 1.0240000000000005e-06    evaluation reward: 9.17
episode: 2373    score: 13.0    memory length: 674484    epsilon: 0.009998020
08555413    steps: 534    lr: 1.0240000000000005e-06    evaluation reward: 9.2
episode: 2374    score: 4.0    memory length: 674763    epsilon: 0.0099980200
08555413    steps: 279    lr: 1.0240000000000005e-06    evaluation reward
: 9.19
episode: 2375    score: 10.0    memory length: 675279    epsilon: 0.009998020
08555413    steps: 516    lr: 1.0240000000000005e-06    evaluation reward
d: 9.21
episode: 2376    score: 7.0    memory length: 675671    epsilon: 0.0099980200
08555413    steps: 392    lr: 1.0240000000000005e-06    evaluation reward
: 9.22
episode: 2377    score: 12.0    memory length: 676234    epsilon: 0.009998020
08555413    steps: 563    lr: 1.0240000000000005e-06    evaluation reward
d: 9.26
episode: 2378    score: 14.0    memory length: 676753    epsilon: 0.009998020
08555413    steps: 519    lr: 1.0240000000000005e-06    evaluation reward
d: 9.3
episode: 2379    score: 13.0    memory length: 677243    epsilon: 0.009998020
08555413    steps: 490    lr: 1.0240000000000005e-06    evaluation reward
d: 9.33
episode: 2380    score: 10.0    memory length: 677767    epsilon: 0.009998020
08555413    steps: 524    lr: 1.0240000000000005e-06    evaluation reward
d: 9.35
episode: 2381    score: 13.0    memory length: 678215    epsilon: 0.009998020
08555413    steps: 448    lr: 1.0240000000000005e-06    evaluation reward
d: 9.35
episode: 2382    score: 7.0    memory length: 678629    epsilon: 0.0099980200
08555413    steps: 414    lr: 1.0240000000000005e-06    evaluation reward
: 9.34
episode: 2383    score: 12.0    memory length: 679257    epsilon: 0.009998020
08555413    steps: 628    lr: 1.0240000000000005e-06    evaluation reward
d: 9.38
episode: 2384    score: 10.0    memory length: 679782    epsilon: 0.009998020
08555413    steps: 525    lr: 1.0240000000000005e-06    evaluation reward
d: 9.38
episode: 2385    score: 9.0    memory length: 680213    epsilon: 0.0099980200
08555413    steps: 431    lr: 1.0240000000000005e-06    evaluation reward
: 9.4
episode: 2386    score: 8.0    memory length: 680648    epsilon: 0.0099980200
08555413    steps: 435    lr: 1.0240000000000005e-06    evaluation reward
: 9.37
```

episode: 2387 score: 13.0 memory length: 681138 epsilon: 0.009998020
008555413 steps: 490 lr: 1.0240000000000005e-06 evaluation reward: 9.41

episode: 2388 score: 11.0 memory length: 681649 epsilon: 0.009998020
008555413 steps: 511 lr: 1.0240000000000005e-06 evaluation reward: 9.45

episode: 2389 score: 7.0 memory length: 682033 epsilon: 0.0099980200
08555413 steps: 384 lr: 1.0240000000000005e-06 evaluation reward : 9.46

episode: 2390 score: 10.0 memory length: 682559 epsilon: 0.009998020
008555413 steps: 526 lr: 1.0240000000000005e-06 evaluation reward: 9.49

episode: 2391 score: 13.0 memory length: 683121 epsilon: 0.009998020
008555413 steps: 562 lr: 1.0240000000000005e-06 evaluation reward: 9.56

episode: 2392 score: 6.0 memory length: 683425 epsilon: 0.0099980200
08555413 steps: 304 lr: 1.0240000000000005e-06 evaluation reward : 9.55

episode: 2393 score: 7.0 memory length: 683800 epsilon: 0.0099980200
08555413 steps: 375 lr: 1.0240000000000005e-06 evaluation reward : 9.52

episode: 2394 score: 6.0 memory length: 684140 epsilon: 0.0099980200
08555413 steps: 340 lr: 1.0240000000000005e-06 evaluation reward : 9.51

episode: 2395 score: 9.0 memory length: 684618 epsilon: 0.0099980200
08555413 steps: 478 lr: 1.0240000000000005e-06 evaluation reward : 9.51

episode: 2396 score: 14.0 memory length: 685133 epsilon: 0.009998020
008555413 steps: 515 lr: 1.0240000000000005e-06 evaluation reward: 9.49

episode: 2397 score: 10.0 memory length: 685623 epsilon: 0.009998020
008555413 steps: 490 lr: 1.0240000000000005e-06 evaluation reward: 9.43

episode: 2398 score: 6.0 memory length: 685976 epsilon: 0.0099980200
08555413 steps: 353 lr: 1.0240000000000005e-06 evaluation reward : 9.4

episode: 2399 score: 9.0 memory length: 686450 epsilon: 0.0099980200
08555413 steps: 474 lr: 1.0240000000000005e-06 evaluation reward : 9.43

episode: 2400 score: 12.0 memory length: 686983 epsilon: 0.009998020
008555413 steps: 533 lr: 1.0240000000000005e-06 evaluation reward: 9.44

episode: 2401 score: 15.0 memory length: 687576 epsilon: 0.009998020
008555413 steps: 593 lr: 1.0240000000000005e-06 evaluation reward: 9.44

episode: 2402 score: 10.0 memory length: 688066 epsilon: 0.009998020
008555413 steps: 490 lr: 1.0240000000000005e-06 evaluation reward: 9.47

episode: 2403 score: 9.0 memory length: 688557 epsilon: 0.0099980200
08555413 steps: 491 lr: 1.0240000000000005e-06 evaluation reward : 9.51

episode: 2404 score: 10.0 memory length: 689077 epsilon: 0.009998020
008555413 steps: 520 lr: 1.0240000000000005e-06 evaluation reward

d: 9.49
episode: 2405 score: 8.0 memory length: 689556 epsilon: 0.0099980200
08555413 steps: 479 lr: 1.0240000000000005e-06 evaluation reward
: 9.5
episode: 2406 score: 6.0 memory length: 689909 epsilon: 0.0099980200
08555413 steps: 353 lr: 1.0240000000000005e-06 evaluation reward
: 9.45
episode: 2407 score: 9.0 memory length: 690366 epsilon: 0.0099980200
08555413 steps: 457 lr: 1.0240000000000005e-06 evaluation reward
: 9.42
episode: 2408 score: 11.0 memory length: 690896 epsilon: 0.0099980200
08555413 steps: 530 lr: 1.0240000000000005e-06 evaluation reward
d: 9.46
episode: 2409 score: 10.0 memory length: 691415 epsilon: 0.0099980200
08555413 steps: 519 lr: 1.0240000000000005e-06 evaluation reward
d: 9.45
episode: 2410 score: 6.0 memory length: 691771 epsilon: 0.0099980200
08555413 steps: 356 lr: 1.0240000000000005e-06 evaluation reward
: 9.45
episode: 2411 score: 9.0 memory length: 692256 epsilon: 0.0099980200
08555413 steps: 485 lr: 1.0240000000000005e-06 evaluation reward
: 9.44
episode: 2412 score: 10.0 memory length: 692764 epsilon: 0.0099980200
08555413 steps: 508 lr: 1.0240000000000005e-06 evaluation reward
d: 9.47
episode: 2413 score: 9.0 memory length: 693220 epsilon: 0.0099980200
08555413 steps: 456 lr: 1.0240000000000005e-06 evaluation reward
: 9.46
episode: 2414 score: 8.0 memory length: 693645 epsilon: 0.0099980200
08555413 steps: 425 lr: 1.0240000000000005e-06 evaluation reward
: 9.46
episode: 2415 score: 7.0 memory length: 694022 epsilon: 0.0099980200
08555413 steps: 377 lr: 1.0240000000000005e-06 evaluation reward
: 9.46
episode: 2416 score: 6.0 memory length: 694364 epsilon: 0.0099980200
08555413 steps: 342 lr: 1.0240000000000005e-06 evaluation reward
: 9.45
episode: 2417 score: 9.0 memory length: 694791 epsilon: 0.0099980200
08555413 steps: 427 lr: 1.0240000000000005e-06 evaluation reward
: 9.37
episode: 2418 score: 5.0 memory length: 695119 epsilon: 0.0099980200
08555413 steps: 328 lr: 1.0240000000000005e-06 evaluation reward
: 9.3
episode: 2419 score: 8.0 memory length: 695551 epsilon: 0.0099980200
08555413 steps: 432 lr: 1.0240000000000005e-06 evaluation reward
: 9.32
episode: 2420 score: 6.0 memory length: 695906 epsilon: 0.0099980200
08555413 steps: 355 lr: 1.0240000000000005e-06 evaluation reward
: 9.24
episode: 2421 score: 5.0 memory length: 696218 epsilon: 0.0099980200
08555413 steps: 312 lr: 1.0240000000000005e-06 evaluation reward
: 9.16
episode: 2422 score: 8.0 memory length: 696660 epsilon: 0.0099980200

```
08555413      steps: 442      lr: 1.0240000000000005e-06      evaluation reward
: 9.16
episode: 2423      score: 9.0      memory length: 697090      epsilon: 0.0099980200
08555413      steps: 430      lr: 1.0240000000000005e-06      evaluation reward
: 9.16
episode: 2424      score: 12.0      memory length: 697574      epsilon: 0.009998020
08555413      steps: 484      lr: 1.0240000000000005e-06      evaluation rewar
d: 9.23
episode: 2425      score: 13.0      memory length: 698061      epsilon: 0.009998020
08555413      steps: 487      lr: 1.0240000000000005e-06      evaluation rewar
d: 9.26
episode: 2426      score: 14.0      memory length: 698593      epsilon: 0.009998020
08555413      steps: 532      lr: 1.0240000000000005e-06      evaluation rewar
d: 9.26
episode: 2427      score: 8.0      memory length: 699024      epsilon: 0.0099980200
08555413      steps: 431      lr: 1.0240000000000005e-06      evaluation reward
: 9.25
episode: 2428      score: 10.0      memory length: 699547      epsilon: 0.009998020
08555413      steps: 523      lr: 1.0240000000000005e-06      evaluation rewar
d: 9.2
episode: 2429      score: 8.0      memory length: 699989      epsilon: 0.0099980200
08555413      steps: 442      lr: 1.0240000000000005e-06      evaluation reward
: 9.14
episode: 2430      score: 6.0      memory length: 700332      epsilon: 0.0099980200
08555413      steps: 343      lr: 4.0960000000000023e-07      evaluation reward
: 9.11
episode: 2431      score: 9.0      memory length: 700787      epsilon: 0.0099980200
08555413      steps: 455      lr: 4.0960000000000023e-07      evaluation reward
: 9.12
episode: 2432      score: 9.0      memory length: 701280      epsilon: 0.0099980200
08555413      steps: 493      lr: 4.0960000000000023e-07      evaluation reward
: 9.12
episode: 2433      score: 11.0      memory length: 701810      epsilon: 0.009998020
08555413      steps: 530      lr: 4.0960000000000023e-07      evaluation rewar
d: 9.16
episode: 2434      score: 8.0      memory length: 702252      epsilon: 0.0099980200
08555413      steps: 442      lr: 4.0960000000000023e-07      evaluation reward
: 9.16
episode: 2435      score: 15.0      memory length: 702823      epsilon: 0.009998020
08555413      steps: 571      lr: 4.0960000000000023e-07      evaluation rewar
d: 9.21
episode: 2436      score: 9.0      memory length: 703304      epsilon: 0.0099980200
08555413      steps: 481      lr: 4.0960000000000023e-07      evaluation reward
: 9.17
episode: 2437      score: 5.0      memory length: 703630      epsilon: 0.0099980200
08555413      steps: 326      lr: 4.0960000000000023e-07      evaluation reward
: 9.14
episode: 2438      score: 13.0      memory length: 704157      epsilon: 0.009998020
08555413      steps: 527      lr: 4.0960000000000023e-07      evaluation rewar
d: 9.16
episode: 2439      score: 9.0      memory length: 704612      epsilon: 0.0099980200
08555413      steps: 455      lr: 4.0960000000000023e-07      evaluation reward
: 9.17
```

```
episode: 2440    score: 8.0    memory length: 705020    epsilon: 0.0099980200
08555413    steps: 408    lr: 4.0960000000000023e-07    evaluation reward
: 9.15
episode: 2441    score: 10.0    memory length: 705499    epsilon: 0.0099980200
008555413    steps: 479    lr: 4.0960000000000023e-07    evaluation reward: 9.17
episode: 2442    score: 10.0    memory length: 706045    epsilon: 0.0099980200
008555413    steps: 546    lr: 4.0960000000000023e-07    evaluation reward: 9.22
episode: 2443    score: 8.0    memory length: 706465    epsilon: 0.0099980200
08555413    steps: 420    lr: 4.0960000000000023e-07    evaluation reward
: 9.24
episode: 2444    score: 10.0    memory length: 706944    epsilon: 0.0099980200
008555413    steps: 479    lr: 4.0960000000000023e-07    evaluation reward: 9.3
episode: 2445    score: 14.0    memory length: 707453    epsilon: 0.0099980200
008555413    steps: 509    lr: 4.0960000000000023e-07    evaluation reward: 9.38
episode: 2446    score: 6.0    memory length: 707851    epsilon: 0.0099980200
08555413    steps: 398    lr: 4.0960000000000023e-07    evaluation reward
: 9.4
episode: 2447    score: 10.0    memory length: 708373    epsilon: 0.0099980200
008555413    steps: 522    lr: 4.0960000000000023e-07    evaluation reward: 9.4
episode: 2448    score: 15.0    memory length: 708920    epsilon: 0.0099980200
008555413    steps: 547    lr: 4.0960000000000023e-07    evaluation reward: 9.47
episode: 2449    score: 9.0    memory length: 709374    epsilon: 0.0099980200
08555413    steps: 454    lr: 4.0960000000000023e-07    evaluation reward
: 9.48
episode: 2450    score: 11.0    memory length: 709795    epsilon: 0.0099980200
008555413    steps: 421    lr: 4.0960000000000023e-07    evaluation reward: 9.44
episode: 2451    score: 9.0    memory length: 710289    epsilon: 0.0099980200
08555413    steps: 494    lr: 4.0960000000000023e-07    evaluation reward
: 9.47
episode: 2452    score: 13.0    memory length: 710865    epsilon: 0.0099980200
008555413    steps: 576    lr: 4.0960000000000023e-07    evaluation reward: 9.51
episode: 2453    score: 9.0    memory length: 711338    epsilon: 0.0099980200
08555413    steps: 473    lr: 4.0960000000000023e-07    evaluation reward
: 9.51
episode: 2454    score: 5.0    memory length: 711650    epsilon: 0.0099980200
08555413    steps: 312    lr: 4.0960000000000023e-07    evaluation reward
: 9.49
episode: 2455    score: 13.0    memory length: 712177    epsilon: 0.0099980200
008555413    steps: 527    lr: 4.0960000000000023e-07    evaluation reward: 9.57
episode: 2456    score: 13.0    memory length: 712704    epsilon: 0.0099980200
008555413    steps: 527    lr: 4.0960000000000023e-07    evaluation reward: 9.62
episode: 2457    score: 10.0    memory length: 713198    epsilon: 0.0099980200
008555413    steps: 494    lr: 4.0960000000000023e-07    evaluation reward
```

d: 9.59
episode: 2458 score: 7.0 memory length: 713604 epsilon: 0.0099980200
08555413 steps: 406 lr: 4.096000000000023e-07 evaluation reward
: 9.55
episode: 2459 score: 9.0 memory length: 714033 epsilon: 0.0099980200
08555413 steps: 429 lr: 4.096000000000023e-07 evaluation reward
: 9.59
episode: 2460 score: 8.0 memory length: 714438 epsilon: 0.0099980200
08555413 steps: 405 lr: 4.096000000000023e-07 evaluation reward
: 9.54
episode: 2461 score: 15.0 memory length: 715011 epsilon: 0.0099980200
08555413 steps: 573 lr: 4.096000000000023e-07 evaluation reward
d: 9.58
episode: 2462 score: 11.0 memory length: 715547 epsilon: 0.0099980200
08555413 steps: 536 lr: 4.096000000000023e-07 evaluation reward
d: 9.59
episode: 2463 score: 7.0 memory length: 715933 epsilon: 0.0099980200
08555413 steps: 386 lr: 4.096000000000023e-07 evaluation reward
: 9.55
episode: 2464 score: 9.0 memory length: 716442 epsilon: 0.0099980200
08555413 steps: 509 lr: 4.096000000000023e-07 evaluation reward
: 9.55
episode: 2465 score: 23.0 memory length: 717224 epsilon: 0.0099980200
08555413 steps: 782 lr: 4.096000000000023e-07 evaluation reward
d: 9.67
episode: 2466 score: 9.0 memory length: 717698 epsilon: 0.0099980200
08555413 steps: 474 lr: 4.096000000000023e-07 evaluation reward
: 9.69
episode: 2467 score: 6.0 memory length: 718094 epsilon: 0.0099980200
08555413 steps: 396 lr: 4.096000000000023e-07 evaluation reward
: 9.7
episode: 2468 score: 11.0 memory length: 718631 epsilon: 0.0099980200
08555413 steps: 537 lr: 4.096000000000023e-07 evaluation reward
d: 9.72
episode: 2469 score: 10.0 memory length: 719052 epsilon: 0.0099980200
08555413 steps: 421 lr: 4.096000000000023e-07 evaluation reward
d: 9.76
episode: 2470 score: 8.0 memory length: 719481 epsilon: 0.0099980200
08555413 steps: 429 lr: 4.096000000000023e-07 evaluation reward
: 9.75
episode: 2471 score: 8.0 memory length: 719958 epsilon: 0.0099980200
08555413 steps: 477 lr: 4.096000000000023e-07 evaluation reward
: 9.72
episode: 2472 score: 8.0 memory length: 720400 epsilon: 0.0099980200
08555413 steps: 442 lr: 4.096000000000023e-07 evaluation reward
: 9.59
episode: 2473 score: 9.0 memory length: 720845 epsilon: 0.0099980200
08555413 steps: 445 lr: 4.096000000000023e-07 evaluation reward
: 9.55
episode: 2474 score: 10.0 memory length: 721322 epsilon: 0.0099980200
08555413 steps: 477 lr: 4.096000000000023e-07 evaluation reward
d: 9.61
episode: 2475 score: 11.0 memory length: 721850 epsilon: 0.0099980200

008555413 steps: 528 lr: 4.0960000000000023e-07 evaluation reward: 9.62
episode: 2476 score: 10.0 memory length: 722345 epsilon: 0.009998020
008555413 steps: 495 lr: 4.0960000000000023e-07 evaluation reward: 9.65
episode: 2477 score: 13.0 memory length: 722980 epsilon: 0.009998020
008555413 steps: 635 lr: 4.0960000000000023e-07 evaluation reward: 9.66
episode: 2478 score: 9.0 memory length: 723473 epsilon: 0.009998020
008555413 steps: 493 lr: 4.0960000000000023e-07 evaluation reward: 9.61
episode: 2479 score: 6.0 memory length: 723829 epsilon: 0.009998020
008555413 steps: 356 lr: 4.0960000000000023e-07 evaluation reward: 9.54
episode: 2480 score: 10.0 memory length: 724331 epsilon: 0.009998020
008555413 steps: 502 lr: 4.0960000000000023e-07 evaluation reward: 9.54
episode: 2481 score: 5.0 memory length: 724642 epsilon: 0.009998020
008555413 steps: 311 lr: 4.0960000000000023e-07 evaluation reward: 9.46
episode: 2482 score: 8.0 memory length: 725077 epsilon: 0.009998020
008555413 steps: 435 lr: 4.0960000000000023e-07 evaluation reward: 9.47
episode: 2483 score: 5.0 memory length: 725383 epsilon: 0.009998020
008555413 steps: 306 lr: 4.0960000000000023e-07 evaluation reward: 9.4
episode: 2484 score: 4.0 memory length: 725662 epsilon: 0.009998020
008555413 steps: 279 lr: 4.0960000000000023e-07 evaluation reward: 9.34
episode: 2485 score: 9.0 memory length: 726119 epsilon: 0.009998020
008555413 steps: 457 lr: 4.0960000000000023e-07 evaluation reward: 9.34
episode: 2486 score: 9.0 memory length: 726596 epsilon: 0.009998020
008555413 steps: 477 lr: 4.0960000000000023e-07 evaluation reward: 9.35
episode: 2487 score: 16.0 memory length: 727217 epsilon: 0.009998020
008555413 steps: 621 lr: 4.0960000000000023e-07 evaluation reward: 9.38
episode: 2488 score: 8.0 memory length: 727649 epsilon: 0.009998020
008555413 steps: 432 lr: 4.0960000000000023e-07 evaluation reward: 9.35
episode: 2489 score: 16.0 memory length: 728233 epsilon: 0.009998020
008555413 steps: 584 lr: 4.0960000000000023e-07 evaluation reward: 9.44
episode: 2490 score: 5.0 memory length: 728545 epsilon: 0.009998020
008555413 steps: 312 lr: 4.0960000000000023e-07 evaluation reward: 9.39
episode: 2491 score: 9.0 memory length: 729028 epsilon: 0.009998020
008555413 steps: 483 lr: 4.0960000000000023e-07 evaluation reward: 9.35
episode: 2492 score: 11.0 memory length: 729558 epsilon: 0.009998020
008555413 steps: 530 lr: 4.0960000000000023e-07 evaluation reward: 9.4

episode: 2493 score: 6.0 memory length: 729947 epsilon: 0.0099980200
08555413 steps: 389 lr: 4.0960000000000023e-07 evaluation reward
: 9.39

episode: 2494 score: 5.0 memory length: 730251 epsilon: 0.0099980200
08555413 steps: 304 lr: 4.0960000000000023e-07 evaluation reward
: 9.38

episode: 2495 score: 9.0 memory length: 730728 epsilon: 0.0099980200
08555413 steps: 477 lr: 4.0960000000000023e-07 evaluation reward
: 9.38

episode: 2496 score: 13.0 memory length: 731315 epsilon: 0.0099980200
08555413 steps: 587 lr: 4.0960000000000023e-07 evaluation reward
: 9.37

episode: 2497 score: 11.0 memory length: 731829 epsilon: 0.0099980200
08555413 steps: 514 lr: 4.0960000000000023e-07 evaluation reward
: 9.38

episode: 2498 score: 8.0 memory length: 732243 epsilon: 0.0099980200
08555413 steps: 414 lr: 4.0960000000000023e-07 evaluation reward
: 9.4

episode: 2499 score: 6.0 memory length: 732586 epsilon: 0.0099980200
08555413 steps: 343 lr: 4.0960000000000023e-07 evaluation reward
: 9.37

episode: 2500 score: 7.0 memory length: 732973 epsilon: 0.0099980200
08555413 steps: 387 lr: 4.0960000000000023e-07 evaluation reward
: 9.32

episode: 2501 score: 8.0 memory length: 733427 epsilon: 0.0099980200
08555413 steps: 454 lr: 4.0960000000000023e-07 evaluation reward
: 9.25

episode: 2502 score: 4.0 memory length: 733700 epsilon: 0.0099980200
08555413 steps: 273 lr: 4.0960000000000023e-07 evaluation reward
: 9.19

episode: 2503 score: 7.0 memory length: 734069 epsilon: 0.0099980200
08555413 steps: 369 lr: 4.0960000000000023e-07 evaluation reward
: 9.17

episode: 2504 score: 7.0 memory length: 734450 epsilon: 0.0099980200
08555413 steps: 381 lr: 4.0960000000000023e-07 evaluation reward
: 9.14

episode: 2505 score: 5.0 memory length: 734758 epsilon: 0.0099980200
08555413 steps: 308 lr: 4.0960000000000023e-07 evaluation reward
: 9.11

episode: 2506 score: 7.0 memory length: 735143 epsilon: 0.0099980200
08555413 steps: 385 lr: 4.0960000000000023e-07 evaluation reward
: 9.12

episode: 2507 score: 14.0 memory length: 735696 epsilon: 0.0099980200
08555413 steps: 553 lr: 4.0960000000000023e-07 evaluation reward
: 9.17

episode: 2508 score: 17.0 memory length: 736257 epsilon: 0.0099980200
08555413 steps: 561 lr: 4.0960000000000023e-07 evaluation reward
: 9.23

episode: 2509 score: 8.0 memory length: 736691 epsilon: 0.0099980200
08555413 steps: 434 lr: 4.0960000000000023e-07 evaluation reward
: 9.21

episode: 2510 score: 8.0 memory length: 737111 epsilon: 0.0099980200
08555413 steps: 420 lr: 4.0960000000000023e-07 evaluation reward

```
: 9.23
episode: 2511  score: 5.0  memory length: 737422  epsilon: 0.0099980200
08555413  steps: 311  lr: 4.0960000000000023e-07  evaluation reward
: 9.19
episode: 2512  score: 7.0  memory length: 737793  epsilon: 0.0099980200
08555413  steps: 371  lr: 4.0960000000000023e-07  evaluation reward
: 9.16
episode: 2513  score: 7.0  memory length: 738199  epsilon: 0.0099980200
08555413  steps: 406  lr: 4.0960000000000023e-07  evaluation reward
: 9.14
episode: 2514  score: 8.0  memory length: 738616  epsilon: 0.0099980200
08555413  steps: 417  lr: 4.0960000000000023e-07  evaluation reward
: 9.14
episode: 2515  score: 10.0  memory length: 739192  epsilon: 0.0099980200
08555413  steps: 576  lr: 4.0960000000000023e-07  evaluation reward
d: 9.17
episode: 2516  score: 7.0  memory length: 739581  epsilon: 0.0099980200
08555413  steps: 389  lr: 4.0960000000000023e-07  evaluation reward
: 9.18
episode: 2517  score: 9.0  memory length: 740034  epsilon: 0.0099980200
08555413  steps: 453  lr: 4.0960000000000023e-07  evaluation reward
: 9.18
episode: 2518  score: 12.0  memory length: 740602  epsilon: 0.0099980200
08555413  steps: 568  lr: 4.0960000000000023e-07  evaluation reward
d: 9.25
episode: 2519  score: 11.0  memory length: 741096  epsilon: 0.0099980200
08555413  steps: 494  lr: 4.0960000000000023e-07  evaluation reward
d: 9.28
episode: 2520  score: 7.0  memory length: 741487  epsilon: 0.0099980200
08555413  steps: 391  lr: 4.0960000000000023e-07  evaluation reward
: 9.29
episode: 2521  score: 6.0  memory length: 741838  epsilon: 0.0099980200
08555413  steps: 351  lr: 4.0960000000000023e-07  evaluation reward
: 9.3
episode: 2522  score: 16.0  memory length: 742543  epsilon: 0.0099980200
08555413  steps: 705  lr: 4.0960000000000023e-07  evaluation reward
d: 9.38
episode: 2523  score: 9.0  memory length: 743017  epsilon: 0.0099980200
08555413  steps: 474  lr: 4.0960000000000023e-07  evaluation reward
: 9.38
episode: 2524  score: 12.0  memory length: 743514  epsilon: 0.0099980200
08555413  steps: 497  lr: 4.0960000000000023e-07  evaluation reward
d: 9.38
episode: 2525  score: 11.0  memory length: 744110  epsilon: 0.0099980200
08555413  steps: 596  lr: 4.0960000000000023e-07  evaluation reward
d: 9.36
episode: 2526  score: 8.0  memory length: 744560  epsilon: 0.0099980200
08555413  steps: 450  lr: 4.0960000000000023e-07  evaluation reward
: 9.3
episode: 2527  score: 9.0  memory length: 745016  epsilon: 0.0099980200
08555413  steps: 456  lr: 4.0960000000000023e-07  evaluation reward
: 9.31
episode: 2528  score: 11.0  memory length: 745455  epsilon: 0.0099980200
```

008555413 steps: 439 lr: 4.0960000000000023e-07 evaluation reward: 9.32
episode: 2529 score: 9.0 memory length: 745930 epsilon: 0.0099980200
008555413 steps: 475 lr: 4.0960000000000023e-07 evaluation reward: 9.33
episode: 2530 score: 17.0 memory length: 746397 epsilon: 0.0099980200
008555413 steps: 467 lr: 4.0960000000000023e-07 evaluation reward: 9.44
episode: 2531 score: 8.0 memory length: 746852 epsilon: 0.0099980200
008555413 steps: 455 lr: 4.0960000000000023e-07 evaluation reward: 9.43
episode: 2532 score: 12.0 memory length: 747474 epsilon: 0.0099980200
008555413 steps: 622 lr: 4.0960000000000023e-07 evaluation reward: 9.46
episode: 2533 score: 8.0 memory length: 747903 epsilon: 0.0099980200
008555413 steps: 429 lr: 4.0960000000000023e-07 evaluation reward: 9.43
episode: 2534 score: 12.0 memory length: 748341 epsilon: 0.0099980200
008555413 steps: 438 lr: 4.0960000000000023e-07 evaluation reward: 9.47
episode: 2535 score: 9.0 memory length: 748859 epsilon: 0.0099980200
008555413 steps: 518 lr: 4.0960000000000023e-07 evaluation reward: 9.41
episode: 2536 score: 11.0 memory length: 749342 epsilon: 0.0099980200
008555413 steps: 483 lr: 4.0960000000000023e-07 evaluation reward: 9.43
episode: 2537 score: 9.0 memory length: 749781 epsilon: 0.0099980200
008555413 steps: 439 lr: 4.0960000000000023e-07 evaluation reward: 9.47
episode: 2538 score: 8.0 memory length: 750186 epsilon: 0.0099980200
008555413 steps: 405 lr: 4.0960000000000023e-07 evaluation reward: 9.42
episode: 2539 score: 6.0 memory length: 750521 epsilon: 0.0099980200
008555413 steps: 335 lr: 4.0960000000000023e-07 evaluation reward: 9.39
episode: 2540 score: 11.0 memory length: 751096 epsilon: 0.0099980200
008555413 steps: 575 lr: 4.0960000000000023e-07 evaluation reward: 9.42
episode: 2541 score: 12.0 memory length: 751659 epsilon: 0.0099980200
008555413 steps: 563 lr: 4.0960000000000023e-07 evaluation reward: 9.44
episode: 2542 score: 9.0 memory length: 752133 epsilon: 0.0099980200
008555413 steps: 474 lr: 4.0960000000000023e-07 evaluation reward: 9.43
episode: 2543 score: 6.0 memory length: 752508 epsilon: 0.0099980200
008555413 steps: 375 lr: 4.0960000000000023e-07 evaluation reward: 9.41
episode: 2544 score: 10.0 memory length: 753026 epsilon: 0.0099980200
008555413 steps: 518 lr: 4.0960000000000023e-07 evaluation reward: 9.41
episode: 2545 score: 6.0 memory length: 753404 epsilon: 0.0099980200
008555413 steps: 378 lr: 4.0960000000000023e-07 evaluation reward: 9.33

episode: 2546 score: 16.0 memory length: 753955 epsilon: 0.009998020
008555413 steps: 551 lr: 4.0960000000000023e-07 evaluation reward: 9.43

episode: 2547 score: 10.0 memory length: 754420 epsilon: 0.009998020
008555413 steps: 465 lr: 4.0960000000000023e-07 evaluation reward: 9.43

episode: 2548 score: 9.0 memory length: 754861 epsilon: 0.0099980200
08555413 steps: 441 lr: 4.0960000000000023e-07 evaluation reward: 9.37

episode: 2549 score: 9.0 memory length: 755285 epsilon: 0.0099980200
08555413 steps: 424 lr: 4.0960000000000023e-07 evaluation reward: 9.37

episode: 2550 score: 12.0 memory length: 755817 epsilon: 0.009998020
008555413 steps: 532 lr: 4.0960000000000023e-07 evaluation reward: 9.38

episode: 2551 score: 8.0 memory length: 756203 epsilon: 0.0099980200
08555413 steps: 386 lr: 4.0960000000000023e-07 evaluation reward: 9.37

episode: 2552 score: 5.0 memory length: 756534 epsilon: 0.0099980200
08555413 steps: 331 lr: 4.0960000000000023e-07 evaluation reward: 9.29

episode: 2553 score: 8.0 memory length: 756940 epsilon: 0.0099980200
08555413 steps: 406 lr: 4.0960000000000023e-07 evaluation reward: 9.28

episode: 2554 score: 8.0 memory length: 757347 epsilon: 0.0099980200
08555413 steps: 407 lr: 4.0960000000000023e-07 evaluation reward: 9.31

episode: 2555 score: 7.0 memory length: 757720 epsilon: 0.0099980200
08555413 steps: 373 lr: 4.0960000000000023e-07 evaluation reward: 9.25

episode: 2556 score: 6.0 memory length: 758078 epsilon: 0.0099980200
08555413 steps: 358 lr: 4.0960000000000023e-07 evaluation reward: 9.18

episode: 2557 score: 15.0 memory length: 758624 epsilon: 0.009998020
008555413 steps: 546 lr: 4.0960000000000023e-07 evaluation reward: 9.23

episode: 2558 score: 24.0 memory length: 759264 epsilon: 0.009998020
008555413 steps: 640 lr: 4.0960000000000023e-07 evaluation reward: 9.4

episode: 2559 score: 12.0 memory length: 759787 epsilon: 0.009998020
008555413 steps: 523 lr: 4.0960000000000023e-07 evaluation reward: 9.43

episode: 2560 score: 12.0 memory length: 760216 epsilon: 0.009998020
008555413 steps: 429 lr: 4.0960000000000023e-07 evaluation reward: 9.47

episode: 2561 score: 7.0 memory length: 760604 epsilon: 0.0099980200
08555413 steps: 388 lr: 4.0960000000000023e-07 evaluation reward: 9.39

episode: 2562 score: 11.0 memory length: 761124 epsilon: 0.009998020
008555413 steps: 520 lr: 4.0960000000000023e-07 evaluation reward: 9.39

episode: 2563 score: 10.0 memory length: 761669 epsilon: 0.009998020
008555413 steps: 545 lr: 4.0960000000000023e-07 evaluation reward:

d: 9.42
episode: 2564 score: 11.0 memory length: 762212 epsilon: 0.009998020
008555413 steps: 543 lr: 4.0960000000000023e-07 evaluation reward
d: 9.44
episode: 2565 score: 6.0 memory length: 762544 epsilon: 0.0099980200
08555413 steps: 332 lr: 4.0960000000000023e-07 evaluation reward
: 9.27
episode: 2566 score: 6.0 memory length: 762876 epsilon: 0.0099980200
08555413 steps: 332 lr: 4.0960000000000023e-07 evaluation reward
: 9.24
episode: 2567 score: 15.0 memory length: 763462 epsilon: 0.009998020
008555413 steps: 586 lr: 4.0960000000000023e-07 evaluation reward
d: 9.33
episode: 2568 score: 4.0 memory length: 763741 epsilon: 0.0099980200
08555413 steps: 279 lr: 4.0960000000000023e-07 evaluation reward
: 9.26
episode: 2569 score: 8.0 memory length: 764101 epsilon: 0.0099980200
08555413 steps: 360 lr: 4.0960000000000023e-07 evaluation reward
: 9.24
episode: 2570 score: 17.0 memory length: 764711 epsilon: 0.009998020
008555413 steps: 610 lr: 4.0960000000000023e-07 evaluation reward
d: 9.33
episode: 2571 score: 13.0 memory length: 765188 epsilon: 0.009998020
008555413 steps: 477 lr: 4.0960000000000023e-07 evaluation reward
d: 9.38
episode: 2572 score: 8.0 memory length: 765619 epsilon: 0.0099980200
08555413 steps: 431 lr: 4.0960000000000023e-07 evaluation reward
: 9.38
episode: 2573 score: 7.0 memory length: 766011 epsilon: 0.0099980200
08555413 steps: 392 lr: 4.0960000000000023e-07 evaluation reward
: 9.36
episode: 2574 score: 7.0 memory length: 766422 epsilon: 0.0099980200
08555413 steps: 411 lr: 4.0960000000000023e-07 evaluation reward
: 9.33
episode: 2575 score: 9.0 memory length: 766877 epsilon: 0.0099980200
08555413 steps: 455 lr: 4.0960000000000023e-07 evaluation reward
: 9.31
episode: 2576 score: 7.0 memory length: 767268 epsilon: 0.0099980200
08555413 steps: 391 lr: 4.0960000000000023e-07 evaluation reward
: 9.28
episode: 2577 score: 9.0 memory length: 767710 epsilon: 0.0099980200
08555413 steps: 442 lr: 4.0960000000000023e-07 evaluation reward
: 9.24
episode: 2578 score: 12.0 memory length: 768267 epsilon: 0.009998020
008555413 steps: 557 lr: 4.0960000000000023e-07 evaluation reward
d: 9.27
episode: 2579 score: 16.0 memory length: 769011 epsilon: 0.009998020
008555413 steps: 744 lr: 4.0960000000000023e-07 evaluation reward
d: 9.37
episode: 2580 score: 5.0 memory length: 769339 epsilon: 0.0099980200
08555413 steps: 328 lr: 4.0960000000000023e-07 evaluation reward
: 9.32
episode: 2581 score: 5.0 memory length: 769667 epsilon: 0.0099980200

```
08555413      steps: 328      lr: 4.0960000000000023e-07      evaluation reward
: 9.32
episode: 2582  score: 11.0    memory length: 770217  epsilon: 0.009998020
08555413      steps: 550      lr: 4.0960000000000023e-07      evaluation reward
d: 9.35
episode: 2583  score: 7.0     memory length: 770592  epsilon: 0.0099980200
08555413      steps: 375      lr: 4.0960000000000023e-07      evaluation reward
: 9.37
episode: 2584  score: 9.0     memory length: 771095  epsilon: 0.0099980200
08555413      steps: 503      lr: 4.0960000000000023e-07      evaluation reward
: 9.42
episode: 2585  score: 10.0    memory length: 771609  epsilon: 0.009998020
08555413      steps: 514      lr: 4.0960000000000023e-07      evaluation reward
d: 9.43
episode: 2586  score: 15.0    memory length: 772317  epsilon: 0.009998020
08555413      steps: 708      lr: 4.0960000000000023e-07      evaluation reward
d: 9.49
episode: 2587  score: 11.0    memory length: 772851  epsilon: 0.009998020
08555413      steps: 534      lr: 4.0960000000000023e-07      evaluation reward
d: 9.44
episode: 2588  score: 15.0    memory length: 773354  epsilon: 0.009998020
08555413      steps: 503      lr: 4.0960000000000023e-07      evaluation reward
d: 9.51
episode: 2589  score: 24.0    memory length: 774001  epsilon: 0.009998020
08555413      steps: 647      lr: 4.0960000000000023e-07      evaluation reward
d: 9.59
episode: 2590  score: 10.0    memory length: 774473  epsilon: 0.009998020
08555413      steps: 472      lr: 4.0960000000000023e-07      evaluation reward
d: 9.64
episode: 2591  score: 13.0    memory length: 774950  epsilon: 0.009998020
08555413      steps: 477      lr: 4.0960000000000023e-07      evaluation reward
d: 9.68
episode: 2592  score: 8.0     memory length: 775392  epsilon: 0.0099980200
08555413      steps: 442      lr: 4.0960000000000023e-07      evaluation reward
: 9.65
episode: 2593  score: 8.0     memory length: 775821  epsilon: 0.0099980200
08555413      steps: 429      lr: 4.0960000000000023e-07      evaluation reward
: 9.67
episode: 2594  score: 13.0    memory length: 776333  epsilon: 0.009998020
08555413      steps: 512      lr: 4.0960000000000023e-07      evaluation reward
d: 9.75
episode: 2595  score: 8.0     memory length: 776728  epsilon: 0.0099980200
08555413      steps: 395      lr: 4.0960000000000023e-07      evaluation reward
: 9.74
episode: 2596  score: 10.0    memory length: 777149  epsilon: 0.009998020
08555413      steps: 421      lr: 4.0960000000000023e-07      evaluation reward
d: 9.71
episode: 2597  score: 8.0     memory length: 777603  epsilon: 0.0099980200
08555413      steps: 454      lr: 4.0960000000000023e-07      evaluation reward
: 9.68
episode: 2598  score: 6.0     memory length: 777959  epsilon: 0.0099980200
08555413      steps: 356      lr: 4.0960000000000023e-07      evaluation reward
: 9.66
```

episode: 2599 score: 7.0 memory length: 778345 epsilon: 0.0099980200
08555413 steps: 386 lr: 4.0960000000000023e-07 evaluation reward
: 9.67
episode: 2600 score: 13.0 memory length: 778827 epsilon: 0.0099980200
08555413 steps: 482 lr: 4.0960000000000023e-07 evaluation reward:
9.73
episode: 2601 score: 11.0 memory length: 779355 epsilon: 0.0099980200
08555413 steps: 528 lr: 4.0960000000000023e-07 evaluation reward:
9.76
episode: 2602 score: 10.0 memory length: 779885 epsilon: 0.0099980200
08555413 steps: 530 lr: 4.0960000000000023e-07 evaluation reward:
9.82
episode: 2603 score: 16.0 memory length: 780602 epsilon: 0.0099980200
08555413 steps: 717 lr: 4.0960000000000023e-07 evaluation reward:
9.91
episode: 2604 score: 7.0 memory length: 780987 epsilon: 0.0099980200
08555413 steps: 385 lr: 4.0960000000000023e-07 evaluation reward
: 9.91
episode: 2605 score: 11.0 memory length: 781505 epsilon: 0.0099980200
08555413 steps: 518 lr: 4.0960000000000023e-07 evaluation reward:
9.97
episode: 2606 score: 12.0 memory length: 782052 epsilon: 0.0099980200
08555413 steps: 547 lr: 4.0960000000000023e-07 evaluation reward:
10.02
episode: 2607 score: 6.0 memory length: 782404 epsilon: 0.0099980200
08555413 steps: 352 lr: 4.0960000000000023e-07 evaluation reward
: 9.94
episode: 2608 score: 21.0 memory length: 782966 epsilon: 0.0099980200
08555413 steps: 562 lr: 4.0960000000000023e-07 evaluation reward:
9.98
episode: 2609 score: 9.0 memory length: 783390 epsilon: 0.0099980200
08555413 steps: 424 lr: 4.0960000000000023e-07 evaluation reward
: 9.99
episode: 2610 score: 7.0 memory length: 783759 epsilon: 0.0099980200
08555413 steps: 369 lr: 4.0960000000000023e-07 evaluation reward
: 9.98
episode: 2611 score: 9.0 memory length: 784285 epsilon: 0.0099980200
08555413 steps: 526 lr: 4.0960000000000023e-07 evaluation reward
: 10.02
episode: 2612 score: 12.0 memory length: 784886 epsilon: 0.0099980200
08555413 steps: 601 lr: 4.0960000000000023e-07 evaluation reward:
10.07
episode: 2613 score: 5.0 memory length: 785179 epsilon: 0.0099980200
08555413 steps: 293 lr: 4.0960000000000023e-07 evaluation reward
: 10.05
episode: 2614 score: 6.0 memory length: 785541 epsilon: 0.0099980200
08555413 steps: 362 lr: 4.0960000000000023e-07 evaluation reward
: 10.03
episode: 2615 score: 9.0 memory length: 785981 epsilon: 0.0099980200
08555413 steps: 440 lr: 4.0960000000000023e-07 evaluation reward
: 10.02
episode: 2616 score: 6.0 memory length: 786308 epsilon: 0.0099980200
08555413 steps: 327 lr: 4.0960000000000023e-07 evaluation reward

```
: 10.01
episode: 2617    score: 5.0    memory length: 786619    epsilon: 0.0099980200
08555413    steps: 311    lr: 4.0960000000000023e-07    evaluation reward
: 9.97
episode: 2618    score: 9.0    memory length: 787088    epsilon: 0.0099980200
08555413    steps: 469    lr: 4.0960000000000023e-07    evaluation reward
: 9.94
episode: 2619    score: 8.0    memory length: 787507    epsilon: 0.0099980200
08555413    steps: 419    lr: 4.0960000000000023e-07    evaluation reward
: 9.91
episode: 2620    score: 15.0    memory length: 788034    epsilon: 0.0099980200
008555413    steps: 527    lr: 4.0960000000000023e-07    evaluation reward
d: 9.99
episode: 2621    score: 8.0    memory length: 788444    epsilon: 0.0099980200
08555413    steps: 410    lr: 4.0960000000000023e-07    evaluation reward
: 10.01
episode: 2622    score: 6.0    memory length: 788773    epsilon: 0.0099980200
08555413    steps: 329    lr: 4.0960000000000023e-07    evaluation reward
: 9.91
episode: 2623    score: 12.0    memory length: 789331    epsilon: 0.0099980200
008555413    steps: 558    lr: 4.0960000000000023e-07    evaluation reward
d: 9.94
episode: 2624    score: 11.0    memory length: 789863    epsilon: 0.0099980200
008555413    steps: 532    lr: 4.0960000000000023e-07    evaluation reward
d: 9.93
episode: 2625    score: 13.0    memory length: 790396    epsilon: 0.0099980200
008555413    steps: 533    lr: 4.0960000000000023e-07    evaluation reward
d: 9.95
episode: 2626    score: 17.0    memory length: 790986    epsilon: 0.0099980200
008555413    steps: 590    lr: 4.0960000000000023e-07    evaluation reward
d: 10.04
episode: 2627    score: 6.0    memory length: 791303    epsilon: 0.0099980200
08555413    steps: 317    lr: 4.0960000000000023e-07    evaluation reward
: 10.01
episode: 2628    score: 6.0    memory length: 791620    epsilon: 0.0099980200
08555413    steps: 317    lr: 4.0960000000000023e-07    evaluation reward
: 9.96
episode: 2629    score: 9.0    memory length: 792114    epsilon: 0.0099980200
08555413    steps: 494    lr: 4.0960000000000023e-07    evaluation reward
: 9.96
episode: 2630    score: 13.0    memory length: 792618    epsilon: 0.0099980200
008555413    steps: 504    lr: 4.0960000000000023e-07    evaluation reward
d: 9.92
episode: 2631    score: 6.0    memory length: 792958    epsilon: 0.0099980200
08555413    steps: 340    lr: 4.0960000000000023e-07    evaluation reward
: 9.9
episode: 2632    score: 7.0    memory length: 793346    epsilon: 0.0099980200
08555413    steps: 388    lr: 4.0960000000000023e-07    evaluation reward
: 9.85
episode: 2633    score: 9.0    memory length: 793801    epsilon: 0.0099980200
08555413    steps: 455    lr: 4.0960000000000023e-07    evaluation reward
: 9.86
episode: 2634    score: 8.0    memory length: 794232    epsilon: 0.0099980200
```



```
08555413    steps: 431    lr: 4.0960000000000023e-07    evaluation reward
: 9.82
episode: 2635    score: 10.0    memory length: 794736    epsilon: 0.009998020
08555413    steps: 504    lr: 4.0960000000000023e-07    evaluation reward: 9.83
episode: 2636    score: 6.0    memory length: 795071    epsilon: 0.0099980200
08555413    steps: 335    lr: 4.0960000000000023e-07    evaluation reward
: 9.78
episode: 2637    score: 10.0    memory length: 795627    epsilon: 0.009998020
08555413    steps: 556    lr: 4.0960000000000023e-07    evaluation reward: 9.79
episode: 2638    score: 5.0    memory length: 795949    epsilon: 0.0099980200
08555413    steps: 322    lr: 4.0960000000000023e-07    evaluation reward
: 9.76
episode: 2639    score: 7.0    memory length: 796340    epsilon: 0.0099980200
08555413    steps: 391    lr: 4.0960000000000023e-07    evaluation reward
: 9.77
episode: 2640    score: 6.0    memory length: 796695    epsilon: 0.0099980200
08555413    steps: 355    lr: 4.0960000000000023e-07    evaluation reward
: 9.72
episode: 2641    score: 14.0    memory length: 797279    epsilon: 0.009998020
08555413    steps: 584    lr: 4.0960000000000023e-07    evaluation reward: 9.74
episode: 2642    score: 5.0    memory length: 797601    epsilon: 0.0099980200
08555413    steps: 322    lr: 4.0960000000000023e-07    evaluation reward
: 9.7
episode: 2643    score: 13.0    memory length: 798129    epsilon: 0.009998020
08555413    steps: 528    lr: 4.0960000000000023e-07    evaluation reward: 9.77
episode: 2644    score: 4.0    memory length: 798405    epsilon: 0.0099980200
08555413    steps: 276    lr: 4.0960000000000023e-07    evaluation reward
: 9.71
episode: 2645    score: 9.0    memory length: 798887    epsilon: 0.0099980200
08555413    steps: 482    lr: 4.0960000000000023e-07    evaluation reward
: 9.74
episode: 2646    score: 12.0    memory length: 799471    epsilon: 0.009998020
08555413    steps: 584    lr: 4.0960000000000023e-07    evaluation reward: 9.7
episode: 2647    score: 8.0    memory length: 799877    epsilon: 0.0099980200
08555413    steps: 406    lr: 4.0960000000000023e-07    evaluation reward
: 9.68
episode: 2648    score: 5.0    memory length: 800199    epsilon: 0.0099980200
08555413    steps: 322    lr: 1.6384000000000001e-07    evaluation reward:
9.64
episode: 2649    score: 7.0    memory length: 800639    epsilon: 0.0099980200
08555413    steps: 440    lr: 1.6384000000000001e-07    evaluation reward:
9.62
episode: 2650    score: 6.0    memory length: 800994    epsilon: 0.0099980200
08555413    steps: 355    lr: 1.6384000000000001e-07    evaluation reward:
9.56
episode: 2651    score: 9.0    memory length: 801462    epsilon: 0.0099980200
08555413    steps: 468    lr: 1.6384000000000001e-07    evaluation reward:
9.57
```

```
episode: 2652  score: 12.0  memory length: 802029  epsilon: 0.009998020
008555413  steps: 567  lr: 1.638400000000001e-07  evaluation reward
: 9.64
episode: 2653  score: 7.0  memory length: 802478  epsilon: 0.0099980200
008555413  steps: 449  lr: 1.638400000000001e-07  evaluation reward:
9.63
episode: 2654  score: 11.0  memory length: 802978  epsilon: 0.009998020
008555413  steps: 500  lr: 1.638400000000001e-07  evaluation reward
: 9.66
episode: 2655  score: 9.0  memory length: 803408  epsilon: 0.0099980200
008555413  steps: 430  lr: 1.638400000000001e-07  evaluation reward:
9.68
episode: 2656  score: 3.0  memory length: 803637  epsilon: 0.0099980200
008555413  steps: 229  lr: 1.638400000000001e-07  evaluation reward:
9.65
episode: 2657  score: 8.0  memory length: 804047  epsilon: 0.0099980200
008555413  steps: 410  lr: 1.638400000000001e-07  evaluation reward:
9.58
episode: 2658  score: 4.0  memory length: 804324  epsilon: 0.0099980200
008555413  steps: 277  lr: 1.638400000000001e-07  evaluation reward:
9.38
episode: 2659  score: 6.0  memory length: 804684  epsilon: 0.0099980200
008555413  steps: 360  lr: 1.638400000000001e-07  evaluation reward:
9.32
episode: 2660  score: 6.0  memory length: 805032  epsilon: 0.0099980200
008555413  steps: 348  lr: 1.638400000000001e-07  evaluation reward:
9.26
episode: 2661  score: 15.0  memory length: 805610  epsilon: 0.009998020
008555413  steps: 578  lr: 1.638400000000001e-07  evaluation reward
: 9.34
episode: 2662  score: 13.0  memory length: 806116  epsilon: 0.009998020
008555413  steps: 506  lr: 1.638400000000001e-07  evaluation reward
: 9.36
episode: 2663  score: 9.0  memory length: 806522  epsilon: 0.0099980200
008555413  steps: 406  lr: 1.638400000000001e-07  evaluation reward:
9.35
episode: 2664  score: 11.0  memory length: 807080  epsilon: 0.009998020
008555413  steps: 558  lr: 1.638400000000001e-07  evaluation reward
: 9.35
episode: 2665  score: 11.0  memory length: 807522  epsilon: 0.009998020
008555413  steps: 442  lr: 1.638400000000001e-07  evaluation reward
: 9.4
episode: 2666  score: 4.0  memory length: 807802  epsilon: 0.0099980200
008555413  steps: 280  lr: 1.638400000000001e-07  evaluation reward:
9.38
episode: 2667  score: 5.0  memory length: 808129  epsilon: 0.0099980200
008555413  steps: 327  lr: 1.638400000000001e-07  evaluation reward:
9.28
episode: 2668  score: 5.0  memory length: 808457  epsilon: 0.0099980200
008555413  steps: 328  lr: 1.638400000000001e-07  evaluation reward:
9.29
episode: 2669  score: 8.0  memory length: 808882  epsilon: 0.0099980200
008555413  steps: 425  lr: 1.638400000000001e-07  evaluation reward:
```

9.29
episode: 2670 score: 6.0 memory length: 809238 epsilon: 0.0099980200
08555413 steps: 356 lr: 1.638400000000001e-07 evaluation reward:
9.18
episode: 2671 score: 10.0 memory length: 809793 epsilon: 0.0099980200
008555413 steps: 555 lr: 1.638400000000001e-07 evaluation reward
: 9.15
episode: 2672 score: 8.0 memory length: 810179 epsilon: 0.0099980200
08555413 steps: 386 lr: 1.638400000000001e-07 evaluation reward:
9.15
episode: 2673 score: 9.0 memory length: 810689 epsilon: 0.0099980200
08555413 steps: 510 lr: 1.638400000000001e-07 evaluation reward:
9.17
episode: 2674 score: 11.0 memory length: 811214 epsilon: 0.0099980200
008555413 steps: 525 lr: 1.638400000000001e-07 evaluation reward
: 9.21
episode: 2675 score: 8.0 memory length: 811677 epsilon: 0.0099980200
08555413 steps: 463 lr: 1.638400000000001e-07 evaluation reward:
9.2
episode: 2676 score: 8.0 memory length: 812097 epsilon: 0.0099980200
08555413 steps: 420 lr: 1.638400000000001e-07 evaluation reward:
9.21
episode: 2677 score: 6.0 memory length: 812451 epsilon: 0.0099980200
08555413 steps: 354 lr: 1.638400000000001e-07 evaluation reward:
9.18
episode: 2678 score: 4.0 memory length: 812731 epsilon: 0.0099980200
08555413 steps: 280 lr: 1.638400000000001e-07 evaluation reward:
9.1
episode: 2679 score: 7.0 memory length: 813121 epsilon: 0.0099980200
08555413 steps: 390 lr: 1.638400000000001e-07 evaluation reward:
9.01
episode: 2680 score: 11.0 memory length: 813648 epsilon: 0.0099980200
008555413 steps: 527 lr: 1.638400000000001e-07 evaluation reward
: 9.07
episode: 2681 score: 6.0 memory length: 814015 epsilon: 0.0099980200
08555413 steps: 367 lr: 1.638400000000001e-07 evaluation reward:
9.08
episode: 2682 score: 6.0 memory length: 814357 epsilon: 0.0099980200
08555413 steps: 342 lr: 1.638400000000001e-07 evaluation reward:
9.03
episode: 2683 score: 3.0 memory length: 814588 epsilon: 0.0099980200
08555413 steps: 231 lr: 1.638400000000001e-07 evaluation reward:
8.99
episode: 2684 score: 4.0 memory length: 814866 epsilon: 0.0099980200
08555413 steps: 278 lr: 1.638400000000001e-07 evaluation reward:
8.94
episode: 2685 score: 5.0 memory length: 815145 epsilon: 0.0099980200
08555413 steps: 279 lr: 1.638400000000001e-07 evaluation reward:
8.89
episode: 2686 score: 9.0 memory length: 815654 epsilon: 0.0099980200
08555413 steps: 509 lr: 1.638400000000001e-07 evaluation reward:
8.83
episode: 2687 score: 6.0 memory length: 816020 epsilon: 0.0099980200

```

08555413      steps: 366      lr: 1.6384000000000001e-07      evaluation reward:
8.78
episode: 2688      score: 12.0      memory length: 816597      epsilon: 0.009998020
008555413      steps: 577      lr: 1.6384000000000001e-07      evaluation reward
: 8.75

```

```

-----
KeyboardInterrupt                                Traceback (most recent call last)
)
Cell In[5], line 56
      54 pylab.ylabel('Rewards')
      55 pylab.title('Episodes vs Reward')
--> 56 pylab.savefig("./save_graph/breakout_dqn.png") # save graph for training visualization
      58 # every episode, plot the play time
      59 print("episode:", e, "  score:", score, "  memory length:",
      60        len(agent.memory), "  epsilon:", agent.epsilon, "  steps:",
step,
      61        "  lr:", agent.optimizer.param_groups[0]['lr'], "
evaluation reward:", np.mean(evaluation_reward))

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/pyplot.py:
1023, in savefig(*args, **kwargs)
      1020 @_copy_docstring_and_deprecators(Figure.savefig)
      1021 def savefig(*args, **kwargs):
      1022     fig = gcf()
-> 1023     res = fig.savefig(*args, **kwargs)
      1024     fig.canvas.draw_idle() # Need this if 'transparent=True', to
reset colors.
      1025     return res

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/figure.py:
3378, in Figure.savefig(self, fname, transparent, **kwargs)
      3374     for ax in self.axes:
      3375         stack.enter_context(
      3376             ax.patch._cm_set(facecolor='none', edgecolor='none'))
-> 3378     self.canvas.print_figure(fname, **kwargs)

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/backend_bases.py:2366, in FigureCanvasBase.print_figure(self, filename, dpi, facecolor, edgecolor, orientation, format, bbox_inches, pad_inches, bbox_extra_artists, backend, **kwargs)
      2362 try:
      2363     # _get_renderer may change the figure dpi (as vector formats
      2364     # force the figure dpi to 72), so we need to set it again here
.
      2365     with cbook._setattr_cm(self.figure, dpi=dpi):
-> 2366         result = print_method(
      2367             filename,
      2368             facecolor=facecolor,
      2369             edgecolor=edgecolor,
      2370             orientation=orientation,
      2371             bbox_inches_restore=_bbox_inches_restore,

```

```

2372         **kwargs)
2373 finally:
2374     if bbox_inches and restore_bbox:

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/backend_bases.py:2232, in FigureCanvasBase._switch_canvas_and_return_print_method.<locals>.<lambda>(*args, **kwargs)

```

2228     optional_kws = { # Passed by print_figure for other renderers
2229         "dpi", "facecolor", "edgecolor", "orientation",
2230         "bbox_inches_restore"}
2231     skip = optional_kws - {*inspect.signature(meth).parameters}
-> 2232     print_method = functools.wraps(meth)(lambda *args, **kwargs: m
meth(
2233         *args, **{k: v for k, v in kwargs.items() if k not in skip
}))
2234 else: # Let third-parties do as they see fit.
2235     print_method = meth

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/backends/backend_agg.py:509, in FigureCanvasAgg.print_png(self, filename_or_obj, metadata, pil_kwargs)

```

462 def print_png(self, filename_or_obj, *, metadata=None, pil_kwargs=
None):
463     """
464     Write the figure to a PNG file.
465     (...)
507     *metadata*, including the default 'Software' key.
508     """
--> 509     self._print_pil(filename_or_obj, "png", pil_kwargs, metadata)

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/backends/backend_agg.py:457, in FigureCanvasAgg._print_pil(self, filename_or_obj, fmt, pil_kwargs, metadata)

```

452 def _print_pil(self, filename_or_obj, fmt, pil_kwargs, metadata=No
ne):
453     """
454     Draw the canvas, then save it using `image.imsave` (to which
455     *pil_kwargs* and *metadata* are forwarded).
456     """
--> 457     FigureCanvasAgg.draw(self)
458     mpl.image.imsave(
459         filename_or_obj, self.buffer_rgba(), format=fmt, origin="u
pper",
460         dpi=self.figure.dpi, metadata=metadata, pil_kwargs=pil_kwa
rgs)

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/backends/backend_agg.py:400, in FigureCanvasAgg.draw(self)

```

396 # Acquire a lock on the shared font cache.
397 with RendererAgg.lock, \
398     (self.toolbar._wait_cursor_for_draw_cm() if self.toolbar

```

```

399         else nullcontext():
--> 400         self.figure.draw(self.renderer)
401         # A GUI class may be need to update a window using this draw,
so
402         # don't forget to call the superclass.
403         super().draw()

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/artist.py:95, in _finalize_rasterization.<locals>.draw_wrapper(artist, renderer, *args, **kwargs)

```

93 @wraps(draw)
94 def draw_wrapper(artist, renderer, *args, **kwargs):
--> 95     result = draw(artist, renderer, *args, **kwargs)
96     if renderer._rasterizing:
97         renderer.stop_rasterizing()

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/artist.py:72, in allow_rasterization.<locals>.draw_wrapper(artist, renderer)

```

69     if artist.get_agg_filter() is not None:
70         renderer.start_filter()
--> 72     return draw(artist, renderer)
73 finally:
74     if artist.get_agg_filter() is not None:

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/figure.py:3175, in Figure.draw(self, renderer)

```

3172         # ValueError can occur when resizing a window.
3174 self.patch.draw(renderer)
-> 3175 mimage._draw_list_compositing_images(
3176     renderer, self, artists, self.suppressComposite)
3178 for sfig in self.subfigs:
3179     sfig.draw(renderer)

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/image.py:131, in _draw_list_compositing_images(renderer, parent, artists, suppress_composite)

```

129 if not_composite or not has_images:
130     for a in artists:
--> 131         a.draw(renderer)
132 else:
133     # Composite any adjacent images together
134     image_group = []

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/artist.py:72, in allow_rasterization.<locals>.draw_wrapper(artist, renderer)

```

69     if artist.get_agg_filter() is not None:
70         renderer.start_filter()
--> 72     return draw(artist, renderer)
73 finally:
74     if artist.get_agg_filter() is not None:

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/axes/_base.py:3064, in _AxesBase.draw(self, renderer)

```

3061 if artists_rasterized:
3062     _draw_rasterized(self.figure, artists_rasterized, renderer)
-> 3064 mimage._draw_list_compositing_images(
3065     renderer, self, artists, self.figure.suppressComposite)
3067 renderer.close_group('axes')
3068 self.stale = False

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/image.py:131, in _draw_list_compositing_images(renderer, parent, artists, suppress_composite)

```

129 if not_composite or not has_images:
130     for a in artists:
--> 131         a.draw(renderer)
132 else:
133     # Composite any adjacent images together
134     image_group = []

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/artist.py:72, in allow_rasterization.<locals>.draw_wrapper(artist, renderer)

```

69 if artist.get_agg_filter() is not None:
70     renderer.start_filter()
---> 72     return draw(artist, renderer)
73 finally:
74     if artist.get_agg_filter() is not None:

```

File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/lines.py:797, in Line2D.draw(self, renderer)

```

794 gc.set_foreground(lc_rgba, isRGBA=True)
796 gc.set_dashes(*self._dash_pattern)
--> 797     renderer.draw_path(gc, tpath, affine.frozen())
798 gc.restore()
800 if self._marker and self._markersize > 0:

```

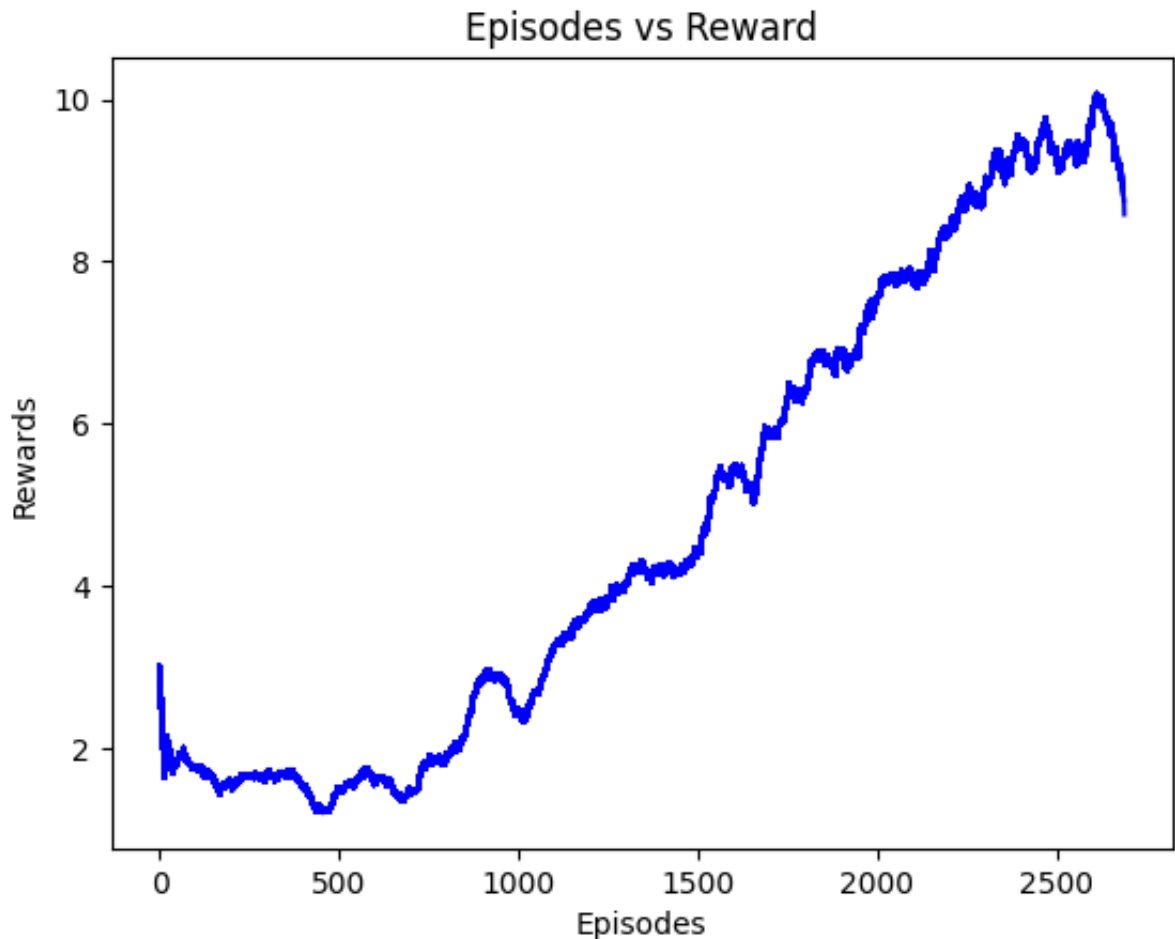
File ~/anaconda3/envs/cv/lib/python3.8/site-packages/matplotlib/backends/backend_agg.py:146, in RendererAgg.draw_path(self, gc, path, transform, rgbFace)

```

144 else:
145     try:
--> 146         self._renderer.draw_path(gc, path, transform, rgbFace)
147     except OverflowError:
148         cant_chunk = ''

```

KeyboardInterrupt:



Visualize Agent Performance

BE AWARE THIS CODE BELOW MAY CRASH THE KERNEL IF YOU RUN THE SAME CELL TWICE.

Please save your model before running this portion of the code.

```
In [ ]: torch.save(agent.policy_net, "./save_model/breakout_dqn_2226.pth")
```

```
In [ ]: import sys
import gym
import torch
import pylab
import random
import numpy as np
from collections import deque
from datetime import datetime
from copy import deepcopy
import torch.nn as nn
import torch.optim as optim
import torch.nn.functional as F
from torch.autograd import Variable
from torch.utils.tensorboard import SummaryWriter
```



```
from utils import find_max_lives, check_live, get_frame, get_init_state
from model import DQN, DQN_LSTM
from config import *

import matplotlib.pyplot as plt
# %load_ext autoreload
# %autoreload 2

from gym.wrappers import RecordVideo # If importing monitor raises issues
import glob
import io
import base64

from IPython.display import HTML
from IPython import display as ipythondisplay

from pyvirtualdisplay import Display

env = gym.make('BreakoutDeterministic-v4', render_mode='rgb_array')
state = env.reset()
number_lives = find_max_lives(env)
state_size = env.observation_space.shape
action_size = 3 #fire, left, and right

# Displaying the game live
def show_state(env, step=0, info=""):
    plt.figure(3)
    plt.clf()
    plt.imshow(env.render(mode='rgb_array'))
    plt.title("%s | Step: %d %s" % ("Agent Playing", step, info))
    plt.axis('off')

    ipythondisplay.clear_output(wait=True)
    ipythondisplay.display(plt.gcf())

# Recording the game and replaying the game afterwards
def show_video():
    mp4list = glob.glob('video/*.mp4')
    if len(mp4list) > 0:
        mp4 = mp4list[0]
        video = io.open(mp4, 'r+b').read()
        encoded = base64.b64encode(video)
        ipythondisplay.display(HTML(data='''<video alt="test" autoplay
            loop controls style="height: 400px;">
            <source src="data:video/mp4;base64,{0}" type="video/mp4"
            </video>'''.format(encoded.decode('ascii'))))
    else:
        print("Could not find video")

def wrap_env(env):
    env = RecordVideo(env, './video')
```

```

    return env

from agent import Agent
action_size = 3

display = Display(visible=0, size=(300, 200))
display.start()

# Load agent
agent = Agent(action_size)
agent.load_policy_net("./save_model/breakout_dqn_2226.pth")
agent.epsilon = 0.0 # Set agent to only exploit the best action

env = wrap_env(env)

done = False
score = 0
step = 0
state, _ = env.reset()
next_state = state
life = number_lives
history = np.zeros([5, 84, 84], dtype=np.uint8)
get_init_state(history, state, HISTORY_SIZE)
frame = 0
while not done:
    # show_state(env, step) # uncommenting this provides another way to vi
    step += 1
    frame += 1

    # Perform a fire action if ball is no longer on screen
    if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
        action = 0
    else:
        action = agent.get_action(np.float32(history[:4, :, :]) / 255.)
    state = next_state

    next_state, reward, done, _, info = env.step(action + 1)

    frame_next_state = get_frame(next_state)
    history[4, :, :] = frame_next_state
    terminal_state = check_live(life, info['lives'])

    life = info['lives']
    r = np.clip(reward, -1, 1)
    r = reward

    # Store the transition in memory
    agent.memory.push(deepcopy(frame_next_state), action, r, terminal_sta
    # Start training after random sample generation
    score += reward

    history[:4, :, :] = history[1:, :, :]
env.close()

```

```
show_video()  
display.stop()
```

```
/home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages/gym/utils/passive_env_checker.py:233: DeprecationWarning: `np.bool8` is a deprecated alias for `np.bool_`. (Deprecated NumPy 1.24)  
    if not isinstance(terminated, (bool, np.bool8)):
```

```
/home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages/gym/wrappers/record_video.py:75: UserWarning: WARN: Overwriting existing videos at /home/wxzhang/assignment5/video folder (try specifying a different `video_folder` for the `RecordVideo` wrapper if this is not desired)  
    logger.warn(  
/home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages/gym/utils/passive_env_checker.py:289: UserWarning: WARN: No render fps was declared in the environment (env.metadata['render_fps'] is None or not defined), rendering may occur at inconsistent fps.  
    logger.warn(  
Moviepy - Building video /home/wxzhang/assignment5/video/rl-video-episode-0.mp4.  
Moviepy - Writing video /home/wxzhang/assignment5/video/rl-video-episode-0.mp4
```

```
Moviepy - Done !  
Moviepy - video ready /home/wxzhang/assignment5/video/rl-video-episode-0.mp4
```

0:00

0:07

```
Out[ ]: <pyvirtualdisplay.display.Display at 0x7f567330d220>
```

video for double DQN

```
In [ ]: torch.save(agent.policy_net, "./save_model/breakout_ddqn_2612.pth")
import sys
import gym
import torch
import pylab
import random
import numpy as np
from collections import deque
from datetime import datetime
from copy import deepcopy
import torch.nn as nn
import torch.optim as optim
import torch.nn.functional as F
from torch.autograd import Variable
from torch.utils.tensorboard import SummaryWriter
from utils import find_max_lives, check_live, get_frame, get_init_state
from model import DQN, DQN_LSTM
from config import *

import matplotlib.pyplot as plt
# %load_ext autoreload
# %autoreload 2

from gym.wrappers import RecordVideo # If importing monitor raises issues
import glob
import io
import base64

from IPython.display import HTML
from IPython import display as ipythondisplay

from pyvirtualdisplay import Display

env = gym.make('BreakoutDeterministic-v4', render_mode='rgb_array')
state = env.reset()
number_lives = find_max_lives(env)
state_size = env.observation_space.shape
action_size = 3 #fire, left, and right

# Displaying the game live
def show_state(env, step=0, info=""):
    plt.figure(3)
    plt.clf()
    plt.imshow(env.render(mode='rgb_array'))
    plt.title("%s | Step: %d %s" % ("Agent Playing", step, info))
    plt.axis('off')

    ipythondisplay.clear_output(wait=True)
    ipythondisplay.display(plt.gcf())
```

```

# Recording the game and replaying the game afterwards
def show_video():
    mp4list = glob.glob('video/*.mp4')
    if len(mp4list) > 0:
        mp4 = mp4list[0]
        video = io.open(mp4, 'r+b').read()
        encoded = base64.b64encode(video)
        ipythondisplay.display(HTML(data='''<video alt="test" autoplay
            loop controls style="height: 400px;">
            <source src="data:video/mp4;base64,{0}" type="video/mp4"
            </video>'''.format(encoded.decode('ascii'))))
    else:
        print("Could not find video")

def wrap_env(env):
    env = RecordVideo(env, './video')
    return env

from agent import Agent
action_size = 3

display = Display(visible=0, size=(300, 200))
display.start()

# Load agent
agent = Agent(action_size)
agent.load_policy_net("./save_model/breakout_dqn_2226.pth")
agent.epsilon = 0.0 # Set agent to only exploit the best action

env = wrap_env(env)

done = False
score = 0
step = 0
state, _ = env.reset()
next_state = state
life = number_lives
history = np.zeros([5, 84, 84], dtype=np.uint8)
get_init_state(history, state, HISTORY_SIZE)
frame = 0
while not done:
    # show_state(env, step) # uncommenting this provides another way to vi
    step += 1
    frame += 1

    # Perform a fire action if ball is no longer on screen
    if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
        action = 0
    else:
        action = agent.get_action(np.float32(history[:4, :, :]) / 255.)
    state = next_state

```

```

next_state, reward, done, _, info = env.step(action + 1)

frame_next_state = get_frame(next_state)
history[4, :, :] = frame_next_state
terminal_state = check_live(life, info['lives'])

life = info['lives']
r = np.clip(reward, -1, 1)
r = reward

# Store the transition in memory
agent.memory.push(deepcopy(frame_next_state), action, r, terminal_state)
# Start training after random sample generation
score += reward

history[:4, :, :] = history[1:, :, :]
env.close()
show_video()
display.stop()

```

```

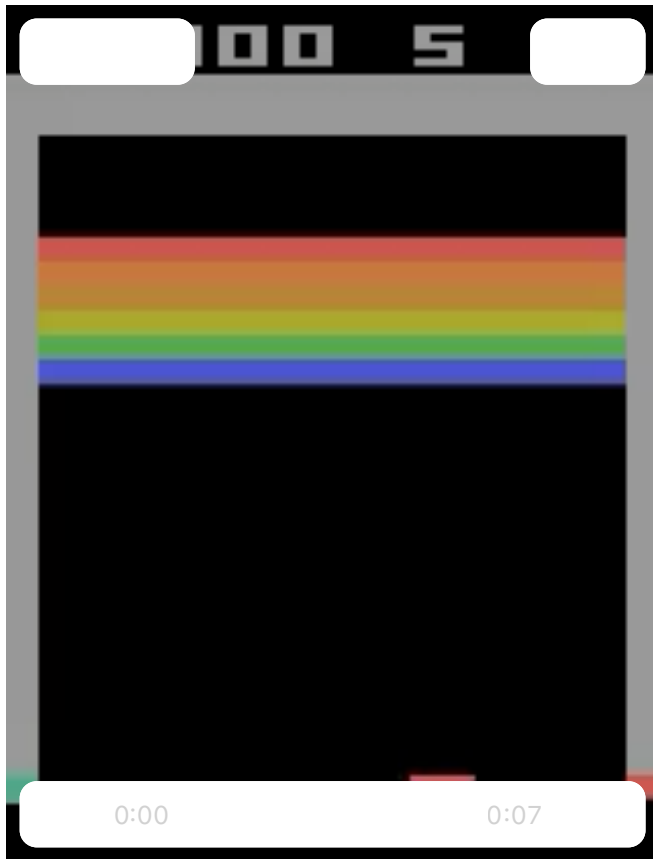
/home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages/gym/utils/passive_env_checker.py:233: DeprecationWarning: `np.bool8` is a deprecated alias for `np.bool_`. (Deprecated NumPy 1.24)
  if not isinstance(terminated, (bool, np.bool8)):
/home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages/gym/wrappers/record_video.py:75: UserWarning: WARN: Overwriting existing videos at /home/wxzhang/assignment5/video folder (try specifying a different `video_folder` for the `RecordVideo` wrapper if this is not desired)
  logger.warn(
/home/wxzhang/anaconda3/envs/cv/lib/python3.8/site-packages/gym/utils/passive_env_checker.py:289: UserWarning: WARN: No render fps was declared in the environment (env.metadata['render_fps'] is None or not defined), rendering may occur at inconsistent fps.
  logger.warn(
Moviepy - Building video /home/wxzhang/assignment5/video/rl-video-episode-0.mp4.
Moviepy - Writing video /home/wxzhang/assignment5/video/rl-video-episode-0.mp4

```

```

Moviepy - Done !
Moviepy - video ready /home/wxzhang/assignment5/video/rl-video-episode-0.mp4

```



Out[]: <pyvirtualdisplay.display.Display at 0x7fc25c2ebaf0>